

Monitoring und Feedback in der stationären Psychosomatik mit dem OQ-Analyst und dem ICD-10-Symptom-Rating

Publikationsgestützte Dissertation

zur Erlangung des akademischen Grades doctor rerum naturalium

(Dr. rer. nat.)

im Fach Psychologie

eingereicht an der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

Dipl.-Psych. Thomas Probst

Präsident der Humboldt-Universität zu Berlin:

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Lebenswissenschaftlichen Fakultät:

Prof. Dr. Richard Lucius

Gutachter:

1. Prof. Dr. T. Fydrich (Humboldt-Universität Berlin)
2. Prof. Dr. M. Ziegler (Humboldt-Universität Berlin)
3. Prof. Dr. A. Mühlberger (Universität Regensburg)

Tag der Verteidigung: 13. März 2015

Selbständigkeitserklärung

Ich versichere, dass ich die Dissertation selbständig und ohne unerlaubte Hilfe angefertigt habe. Die benutzten Hilfsmittel sowie die Literatur sind vollständig angegeben. Ich habe die Dissertation an keiner anderen Universität eingereicht und besitze keinen Doktorgrad im Fach Psychologie. Die Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät II der Humboldt-Universität zu Berlin vom 17.01.2005, zuletzt geändert am 13.02.2006, veröffentlicht im Amtlichen Mitteilungsblatt Nr. 34/2006 ist mir bekannt.

Berlin,

Thomas Probst

Abstract - Deutsch

Diese publikationsgestützte Dissertation umfasst Arbeiten zum Thema Monitoring und Feedback in der stationären Psychosomatik. In zwei Arbeiten wurde untersucht, ob die Effektivität von psycho-somatischen Behandlungen gesteigert werden kann, wenn Therapeuten regelmäßig Feedbackberichte der Monitoring- und Feedbacksoftware OQ-Analyst erhalten. Hier zeigte sich, dass Feedback den Outcome von Patienten mit erhöhtem Risiko einer Verschlechterung verbessern kann. Bei den restlichen Patienten dagegen konnte Feedback den Therapieprozess nur schwach und nur kurzfristig für eine Behandlungswoche verbessern. In einer weiteren Arbeit wurde untersucht, ob sich unterschiedliche Belastungsgrade im wöchentlichen Monitoring differenziell verändern. Hierfür wurde das ICD-10-Symptomrating herangezogen. Die Ergebnisse lieferten Hinweise darauf, dass sich depressive Beschwerden und Angstbeschwerden schon in der ersten Behandlungswoche verbessern, somatoforme Beschwerden jedoch erst im späteren Behandlungsverlauf. Die vierte Arbeit setzte sich mit der Fragestellung auseinander, ob Therapiebeziehung, soziale Unterstützung, Motivation, und Lebensereignisse mit bedeutsamen Verbesserungen und bedeutsamen Verschlechterungen assoziiert sind. Während alle vier Bereiche mit bedeutsamen Verbesserungen verknüpft waren, waren nur soziale Unterstützung und Lebensereignisse mit bedeutsamen Verschlechterungen assoziiert.

Schlagwörter: Monitoring, Feedback, Psychosomatik, Psychotherapieforschung

Abstract - English

This dissertation is based on four publications on monitoring and feedback in psychosomatic in-patient therapy. In two publication it was investigated whether the effectiveness of psychosomatic in-patient therapy can be increased when therapists routinely receive the feedback reports of the monitoring and feedback software OQ-Analyst. Feedback could improve the outcome of patients at risk of treatment failure. For the remaining patients, feedback could optimize the therapy process only slightly for one treatment week. In another publication, the ICD-10-Symptom-Rating was used to explore how different symptoms change on a weekly basis. While depressive symptoms and symptoms of anxiety disorders improved already in the first treatment week, more treatment weeks were required to improve somatoform symptoms. In a fourth publication, associations between considerable positive as well as considerable negative change and the therapeutic alliance, social support, motivation, and life events were investigated. While all four factors were associated with considerable improvements, only social support and life events correlated with considerable deteriorations.

Keywords: Monitoring, feedback, psychosomatics, patient-focused research.

Danksagung

Bei Prof. Dr. Thomas Fydrich bedanke ich mich für die Betreuung dieser Dissertation an der Humboldt-Universität zu Berlin. Prof. Dr. Michael J. Lambert möchte ich herzlich für den beständigen Zuspruch und die menschliche Unterstützung danken. Mein Dank gilt auch den Chefärzten der beteiligten psychosomatischen Kliniken, Prof. Dr. Thomas H. Loew und PD Dr. Reiner W. Dahlbender, und all den Therapeuten, Pflegekräften sowie Patienten dieser Kliniken für die Teilnahme an diesem Projekt. Auch möchte ich mich bei Johanna Karg, Elena Kopp und Carolin Schmalhofer bedanken, die mich während ihres Psychologiestudiums im Rahmen von Praktika bei der Datenerhebung in den Kliniken unterstützt haben. Zu guter Letzt bedanke ich mich ganz herzlich bei Prof. Dr. Karin Tritt, die mir immer zur Seite stand und unablässig darauf erpicht war, das Beste aus mir herauszukitzeln. Für meine Frau Bori!

Inhaltsverzeichnis

I.	Zusammenfassung / Abstract	6
II.	Hintergrund	8
II. 1.	Patient-focused research	8
II. 2.	OQ-Analyst.....	11
II. 3.	ICD-10-Symptom-Rating	20
II. 4.	Forschungsagenda	21
III.	Publikationen.....	24
III. 1.	Publikation 1	24
III. 2.	Publikation 2	33
III. 3.	Publikation 3	42
III. 4.	Publikation 4	60
IV.	Diskussion und Ausblick.....	92
IV. 1.	Publikationen 1 & 2	92
IV. 2.	Publikation 3	96
IV. 3.	Publikation 4	99
IV. 4.	Zusammenfassende Betrachtung.....	101
V.	Literaturverzeichnis zum Klappentext	102

Abbildungsverzeichnis

Abbildung 1: Feedbackbericht, der den Therapieverlauf im „Outcome Questionnaire“ (OQ-45) anzeigt. Mit freundlicher Genehmigung von OQ-Measures.....	13
Abbildung 2: Feedbackbericht zu den Ergebnissen im „Assessment for Signal Cases“ (ASC). Mit freundlicher Genehmigung von OQ-Measures.	17

Abkürzungsverzeichnis

ASC	Assessment for Signal Cases
CFIT	Contextual Feedback Intervention Theory
CG	Kontrollgruppe
CST	Clinical Support Tools
EBM	Evidence Based Medicine
EBP	Evidence Based Practice
EG	Experimentalgruppe=Feedbackgruppe
EST	Empirically Supported Treatments
ISR	ICD-10-Symptom-Rating
NOT	Not-On-Track
OQ-45	Outcome Questionnaire
OT	On-Track
RCI	Reliable Change Index
ROM	Routine Outcome Monitoring
TAU	Treatment-As-Usual

I. Zusammenfassung / Abstract

Hintergrund: Im Rahmen der patientenorientierten Psychotherapieforschung werden kontinuierlich während einer Therapie behandlungsrelevante Kriterien aus Patientensicht erhoben (Monitoring) und dem behandelnden Therapeuten Feedback darüber zur Verfügung gestellt.

Methode: Patienten zweier psychosomatischer Kliniken wurden bei Aufnahme in die Experimentalgruppe (EG) oder in die Kontrollgruppe (CG) randomisiert. Beide Gruppen nahmen am Monitoring teil, indem sie während ihres Klinikaufenthaltes einmal wöchentlich Fragebögen („Outcome Questionnaire“ OQ-45, „Assessment for Signal Cases“ ASC und „ICD-10-Symptom-Rating“ ISR) ausfüllten. Kontinuierliches Feedback erhielten die Therapeuten für ihre Patienten der EG sowohl zum Therapieverlauf im OQ-45 als auch zu den durch den ASC erhobenen Faktoren (Therapiebeziehung, soziale Unterstützung, Motivation, Lebensereignisse). Die Feedbackberichte wurden mit dem OQ-Analyst erstellt: Der Therapieverlauf eines Patienten X im OQ-45 wurde dabei im Vergleich zum erwarteten Therapieverlauf dieses Patienten X rückgemeldet. Der erwartete Therapieverlauf eines Patienten X entsprach dem Therapieverlauf anderer Patienten, die bei Aufnahme einen ähnlich hohen Beschwerdescore im OQ-45 erreichten wie Patient X.

Ergebnisse: Bei Patienten mit negativer Abweichung vom erwarteten Therapieverlauf (mind. eine Standardabweichung) zeigte die EG einen besseren Prä-Post Outcome als die CG auf der „Gesamtskala“ des OQ-45 (Effektstärke=0.54). Dagegen war bei Patienten, die nicht um mind. eine Standardabweichung negativ vom erwarteten Therapieverlauf abwichen, der Therapieprozess der EG nur für eine Behandlungswoche und nur auf der OQ-45 Subskala „Symptombelastung“ geringfügig besser als der Therapieprozess der CG (Effektstärke=0.12). Genauere Analysen der wöchentlichen Therapieverläufe zeigten, dass sich Patienten während der Aufnahmewoche auf den Skalen des OQ-45 und des ISR unterschiedlich belastet einschätzten und sich in den ersten vier Wochen der Behandlung in diesen Skalen differenziell veränderten. Alle vier im ASC erhobenen Faktoren hingen mit bedeutsamen Verbesserungen zusammen. Mit bedeutsamen Verschlechterungen waren jedoch nur die soziale Unterstützung und Lebensereignisse, nicht aber die Therapiebeziehung und die Motivation assoziiert.

Schlussfolgerung: Der OQ-Analyst kann zur Prävention von Verschlechterungen in der stationären Psychosomatik beitragen. Zur Besserung der Zwangs- bzw. Essstörungssymptomatik scheinen mehr als vier Behandlungswochen nötig zu sein. Um bei Patienten mit bedeutsamer Verschlechterung das Behandlungsergebnis zu optimieren, könnten die Etablierung eines sozialen Netzwerks und das Coping von Lebensereignissen wichtig sein.

Abstract

Background: In “patient focused research”, patients are routinely monitored with patient-reported outcome measures during the course of psychotherapy so that continuous feedback on the monitoring data can be provided to therapists.

Method: Patients of two psychosomatic clinics were randomized either into the experimental group (EG) or into the control group (CG). While both groups were monitored with patient-reported outcome measures (“Outcome Questionnaire” OQ-45, “Assessment for Signal Cases” ASC, and “ICD-10-Symptom-Rating” ISR) on a weekly basis during their in-patient stay, continuous feedback on patient progress (OQ-45) and on domains assessed with the ASC (therapeutic alliance, social support, motivation, life events) was provided to therapists only for their in-patients belonging to the EG. The feedback reports provided to the therapists were generated by the monitoring- and feedbacksoftware OQ-Analyst. The feedback on patient progress displayed the individual patient’s progress on the OQ-45 in comparison to an expected recovery curve defined as the progress of patients with scores on the OQ-45 comparable to the individual patient at intake.

Results: For in-patients deviating negatively from their expected recovery curves by at least one standard deviation, the EG reached an improved outcome compared to the CG on the OQ-45 “total scale” ($g=0.54$). For in-patients not deviating negatively from their expected recovery curves by at least one standard deviation, however, the EG surpassed the CG only for one treatment week and merely slightly on the OQ-45 subscale “symptom distress” ($g=0.12$). A more detailed analysis of the weekly assessments revealed that the patients’ impairments during the intake week as well as their weekly patterns of change varied among the OQ-45 and ISR scales. All four domains assessed with the ASC were related to considerable

improvements. The social support scale and the life events scale of the ASC, but neither the ASC therapeutic alliance scale nor the ASC motivation scale were associated with considerable deteriorations.

Conclusion: The OQ-Analyst can be used to prevent treatment failure in psychosomatic in-patient therapy. More than four treatment weeks seem to be necessary to improve symptoms typical for eating disorders and obsessive-compulsive disorders. Building a social network and coping life events could be of importance when trying to optimize treatment for patients with considerable deteriorations.

II. Hintergrund

II. 1. Patient-focused research

Die hier vorgelegte Dissertation ist in der patientenorientierten Psychotherapieforschung („patient focused research“) zu verorten. Die „patient focused research“ wurde von Howard, Moras, Brill, Martinovich, und Lutz (1996) als Ergänzung zur „treatment focused research“ vorgeschlagen. Um die Vorgehensweise beider Forschungsrichtungen aufzuzeigen, wird im Folgenden zuerst auf die „treatment focused research“ eingegangen und danach ausführlicher die „patient focused research“ dargestellt.

Im Rahmen der „treatment focused research“ sollen Behandlungstechniken sowohl unter den Bedingungen randomisierter kontrollierter Studien („efficacy“) als auch unter naturalistischen Routinebedingungen („effectiveness“) evaluiert werden, um „Empirically Supported Treatments“ (EST) zu identifizieren (Chambless, & Hollon, 1998). Die EST sind wesentlicher Bestandteil der triadisch konzeptualisierten „Evidence Based Medicine“ (EBM; Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996) bzw. „Evidence Based Practice“ (EBP; Spring, 2007), im Rahmen derer sie den Strang der „best research evidence“ repräsentieren sollen.

Im Bereich der Psychotherapie wurde der EBM-Strang „best research evidence“ mehr als die anderen beiden EBM-Stränge („clinical expertise“, „patient preferences“) kontrovers diskutiert (z. B. Henningsen, & Rudolf, 2000; Lilienfeld, Ritschel, Lynn, Cautin, & Latzman, 2013; Tschuske, 2005). Als wesentlicher Kritikpunkt wurde angeführt, dass das Konzept der EST zu stark am medizinischen Modell orientiert sei, das darauf abzielt, eine effektive Behandlungstechnik für eine bestimmte Störung zu identifizieren. Dieses Modell sei auf die Psychotherapie nicht eins zu eins übertragbar: Einerseits leiden Psychotherapiepatienten in der Regel meist nicht unter einer bestimmten psychischen Störung, sondern komorbid unter mehreren psychischen Störungen (Tschuske, 2005). Andererseits ist in der Psychotherapie die Behandlungstechnik nur zu einem relativ geringem Prozentsatz (15% - 17%) für den Therapieerfolg verantwortlich (Asay, & Lambert, 1999; Cuijpers et al., 2012).

Im Gegensatz zur „treatment focused research“ steht in der „patient focused research“ nicht die Behandlungstechnik im Vordergrund der Forschung, sondern die Entwicklung des individuellen Patienten während einer Psychotherapie sowie darauf wirkende Einflussfaktoren. Der Therapieverlauf sowie mögliche Einflussfaktoren werden dabei durch psychometrische Fra-

gebögen („patient-reported outcome measures“; Greenhalgh, 2009) erhoben, die die Patienten kontinuierlich (z. B. wöchentlich) während der Psychotherapie bearbeiten. Die in den psychometrischen Fragebögen erfassten behandlungsrelevanten Kriterien (z. B. psychische Beschwerden) werden im Rahmen der „patient focused research“ dem zuständigen Therapeuten rückgemeldet, so dass dieser die Patientensicht zur optimalen Gestaltung der Psychotherapie nutzen kann. Dieses Vorgehen hat Ähnlichkeit zur „personalized medicine“ (Jain, 2002), im Rahmen derer molekulare Informationen vom Patienten erhoben werden und diese Informationen dafür genutzt werden sollen, dem Patienten eine optimale Behandlung zu Gute kommen zu lassen (zur „personalized medicine“ bei der psychopharmakologischen Behandlung von psychischen Störungen siehe z. B. Hall-Flavin et al., 2012; 2013).

Die zwei wesentlichen Merkmale der „patient focused research“ - 1) kontinuierliches Erheben behandlungsrelevanter Kriterien während der Psychotherapie aus Patientensicht und 2) kontinuierliche Rückmeldung dieser Kriterien an den zuständigen Therapeuten - werden im Folgenden genauer dargestellt.

Das kontinuierliche Erheben von behandlungsrelevanten Kriterien während einer Psychotherapie wird auch als Monitoring oder „Routine Outcome Monitoring“ (ROM) bezeichnet (Boswell, Kraus, Miller, & Lambert, 2013). Die durch ROM gewonnenen Daten wurden in der bisherigen Psychotherapieforschung zur Untersuchung verschiedener Fragestellungen verwendet. Im Folgenden werden jene Fragestellungen und dazu vorliegende Ergebnisse zusammengefasst, die für die hier vorgelegte Dissertation von besonderer Relevanz sind.

Durch ROM kann untersucht werden, welche behandlungsrelevanten Kriterien sich wann während einer Psychotherapie verändern. Howard, Lueger, Maling und Martinovich (1993) formulierten hierzu das „phase model of therapeutic change“, nach dem sich in ambulanten Psychotherapien in der ersten Behandlungsphase („remoralization“) v. a. das Wohlbefinden verbessert, in der nächsten Behandlungsphase („remediation“) besonders Symptome reduziert werden und in der folgenden Behandlungsphase („rehabilitation“) die Verbesserung sozialer sowie interpersoneller Probleme im Vordergrund steht (siehe auch Howard et al., 1996). Andere Studien in diesem Zusammenhang zeigten, dass sich unterschiedliche Symptome (Kopta, Howard, Lowry, & Beutler, 1994) sowie unterschiedliche interpersonelle Probleme (Maling, Gurtman, & Howard, 1995) unterschiedlich schnell während einer Psychotherapie verändern.

ROM ermöglicht zudem die Identifikation von bedeutsamen Verbesserungen bzw. bedeutsamen Verschlechterungen eines Patienten während einer laufenden Therapie. Dies ist bzgl. der Prädiktion des weiteren Therapieverlaufs eines Patienten von Relevanz, denn Patienten mit bedeutsamen Verschlechterungen während einer Therapie haben ein erhöhtes Risiko für einen schlechten Therapieoutcome (Lambert, Whipple, Bishop et al., 2002; Spielmans, Masters, & Lambert, 2006) und Patienten mit bedeutsamen Verbesserungen während einer Therapie dagegen eine erhöhte Chance für einen positiven Therapieoutcome (Aderka, Nickerson, Bøe, & Hofmann, 2012).

Auch können die im Rahmen des ROM erhobenen Daten - wie bereits erwähnt - kontinuierlich während der Behandlung an den zuständigen Therapeuten rückgemeldet werden. International wurde der Effekt von solchen Feedbackinterventionen auf den Patientenoutcome in zahlreichen Studien und einigen Meta-Analysen untersucht (siehe Castonguay et al., 2013). Im deutschsprachigen Raum wurden Feedbackeffekte bisher in der ambulanten Psychotherapie (Wittmann et al., 2011), in der stationären Psychiatrie (Puschner, Schöfer, Knaup, & Beck, 2009) und in der stationären Psychosomatik (Bauer, 2004; Berking, Orth, & Lutz, 2006) untersucht. Zudem ist eine weitere Feedbackstudie in der stationären Psychosomatik geplant (Kraft, Koellner, & Lutz, 2008). Die beiden abgeschlossenen Studien zur stationären Psychosomatik werden im Folgenden etwas genauer dargestellt, da sich diese Arbeit auf dieses psychotherapeutische Setting bezieht.

In der Studie von Berking et al. (2006) wurden Patienten einer verhaltenstherapeutisch-orientierten psychosomatischen Klinik auf Kontroll- und Feedbackgruppe randomisiert. Nur für Patienten der Feedbackgruppe erhielt der behandelnde Therapeut wöchentlich Rückmeldung über den Therapieverlauf (T-Werte der Fragebögen und Ziel-Erreichungs-Angaben). Die Feedbackgruppe zeigte einen signifikant besseren Outcome als die Kontrollgruppe.

Bei Bauer (2004) wurden in einer psychoanalytisch-orientierten psychosomatischen Klinik zu unterschiedlichen Zeiten vier Studienphasen realisiert („historical cohort design“): In allen vier Phasen bearbeiteten die Patienten bei Aufnahme, Zwischenmessung (3 - 5 Wochen nach Aufnahme) und bei Entlassung Fragebögen. In Phase 1 erhielten die Behandler über die Fragebogenergebnisse keinerlei Rückmeldungen. In Phase 2 wurden die Fragebogendaten von bereits entlassenen Patienten rückgemeldet. In Phase 3 wurden nicht nur die Daten der entlassenen Patienten rückgemeldet, sondern auch der Therapieverlauf, wie er sich bei noch in Behandlung befindlichen Patienten zwischen Aufnahme und Zwischenmessung in den Fragebögen zeigte. In Phase 4 wurde die Rückmeldung über den Therapieverlauf zwi-

schen Aufnahme und Zwischenmessung noch ergänzt durch eine Behandlungsempfehlung darüber, ob die Behandlung eines Patienten fortgesetzt oder beendet werden sollte. Der Outcome der Patienten aus Phase 1 und Phase 2 wurde verglichen mit dem Outcome der Patienten aus Phase 3 und Phase 4. Wenn Feedback über den Therapieverlauf von Aufnahme zu Zwischenmessung gegeben wurde (Phase 3 und Phase 4) war der Outcome nicht signifikant besser als ohne Feedback (Phase 1 und Phase 2). Dieses Ergebnis zeigte sich sowohl über alle Patienten als auch bei solchen Patienten, die sich zur Zwischenmessung in den Fragebögen als nonfunktional oder funktional einschätzten. Zur Klassifikation als nonfunktional bzw. funktional wurden Cut-Off Werte gemäß dem c-Kriterium von Jacobson und Truax (1991) verwendet.

Im Gegensatz zu diesen beiden bisherigen Studien zu Monitoring und Feedback in der stationären Psychosomatik wurde in dieser Arbeit die im Folgenden näher beschriebene Monitoring- und Feedbacksoftware OQ-Analyst eingesetzt.

II. 2. OQ-Analyst

Der OQ-Analyst (Lambert, 2012) ist eines von mehreren sog. „bona fide measurement systems“ wie Monitoring- und Feedbackprogramme zur Qualitätssicherung und –verbesserung von Psychotherapien auch bezeichnet werden (Castonguay et al., 2013). Andere in dieser Arbeit nicht näher dargestellte, bei Boswell et al. (2013) beschriebene, „bona fide measurement systems“ sind z. B. „Clinical Outcomes in Routine Evaluation“ (CORE), „Treatment Outcome Package“ (TOP), „The Partners for Change Outcome Management System: International Center for Clinical Excellence“ (PCOMS ICCE).

Im OQ-Analyst sind zwei Selbsteinschätzungsfragebögen für Patienten integriert: Der eine Fragebogen ist der „Outcome Questionnaire“ (OQ-45; Lambert, Kahler, Harmon, Burlingame, & Shimokawa, 2011; auch Ergebnisfragebogen EB-45 genannt; Lambert, Hannover, Nisslmüller, Richard, & Kordy, 2002). Der OQ-45 besteht aus einer Gesamtskala sowie drei Subskalen (Symptombelastung, zwischenmenschliche Beziehungen, soziale Integration) und erhebt die psychische, interpersonelle und soziale Beeinträchtigung bezogen auf die vergangene Woche. Die Gesamtskala des OQ-45 wird im OQ-Analyst dafür verwendet, einen Feedbackbericht über den Therapieverlauf zu erstellen. Der andere Fragebogen ist der „Assessment for Signal Cases“ (ASC; Lambert et al., 2007). Der ASC dient zur Erhebung von vier Faktoren, die einen Einfluss auf die Psychotherapie nehmen können: Therapiebeziehung,

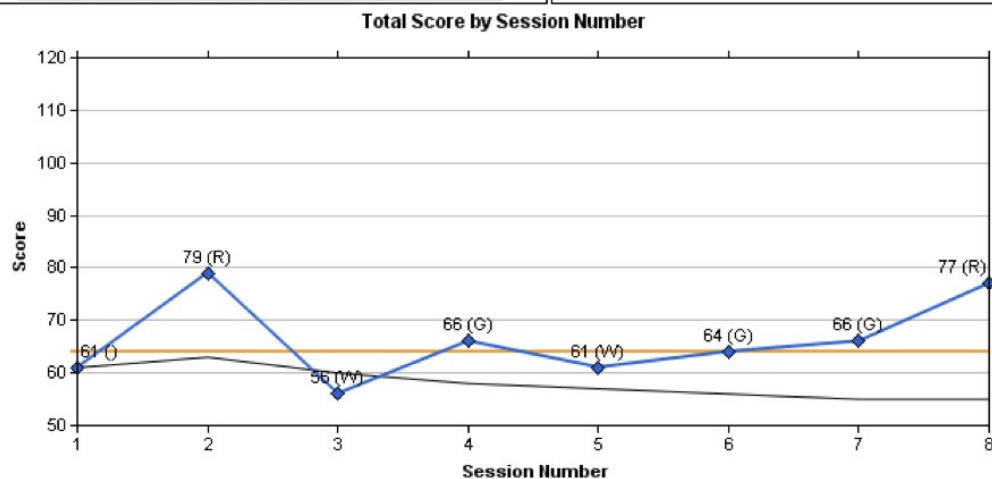
soziale Unterstützung, Motivation und Lebensereignisse. Diese vier Faktoren werden im ASC bezogen auf die vergangene Woche erhoben. Der ASC wird im OQ-Analyst im Rahmen von „Clinical Support Tools“ (CST) verwendet.

Im Folgenden wird zuerst auf den Feedbackbericht zum Therapieverlauf eingegangen. Danach werden die CST des OQ-Analyst dargestellt.

Sind mindestens zwei OQ-45 Messungen eines Patienten in den OQ-Analyst übertragen, so kann im OQ-Analyst ein Feedbackbericht über den Therapieverlauf dieses Patienten erzeugt werden. Abbildung 1 zeigt beispielhaft einen solchen Feedbackbericht.

Name:	ID:	Alert Status:	Red
Session Date:	Session: 8	Most Recent Score:	77
Clinician:	Clinic:	Initial Score:	61
Diagnosis:	Unknown Diagnosis	Change From Initial:	Reliably Worse
Algorithm:	Empirical	Current Distress Level:	Moderate
Instrument:	OQ@-45.2		

Most Recent Critical Item Status:		Subscales	Current	Output. Norm	Comm. Norm
8. Suicide - I have thoughts of ending my life.	Never	Symptom Distress:	53	49	25
11. Substance Abuse - After heavy drinking, I need a drink the next morning to get going.	Never	Interpersonal Relations:	16	20	10
26. Substance Abuse - I feel annoyed by people who criticize my drinking.	Never	Social Role:	8	14	10
32. Substance Abuse - I have trouble at work/school because of drinking or drug use.	Never	Total:	77	83	45
44. Work Violence - I feel angry enough at work/school to do something I might regret.	Never				



Graph Label Legend:

(R) = **Red**: High chance of negative outcome (Y) = **Yellow**: Some chance of negative outcome
(G) = **Green**: Making expected progress (W) = **White**: Functioning in normal range

Feedback Message:

The patient is deviating from the expected response to treatment. They are not on track to realize substantial benefit from treatment. Chances are they may drop out of treatment prematurely or have a negative treatment outcome. Steps should be taken to carefully review this case and identify reasons for poor progress. It is recommended that you be alert to the possible need to improve the therapeutic alliance, reconsider the client's readiness for change and the need to renegotiate the therapeutic contract, intervene to strengthen social supports, or possibly alter your treatment plan by intensifying treatment, shifting intervention strategies, or decide upon a new course of action, such as referral for medication. Continuous monitoring of future progress is highly recommended.

Abbildung 1: Feedbackbericht, der den Therapieverlauf im „Outcome Questionnaire“ (OQ-45) anzeigt. Mit freundlicher Genehmigung von OQ-Measures.

Abbildung 1 zeigt, dass der Feedbackbericht zum Therapieverlauf aus Textfeldern und einem Liniendiagramm besteht. Im Liniendiagramm entspricht die blaue Linie dem Therapieverlauf des individuellen Patienten X, die schwarze Linie zeigt den erwarteten Therapieverlauf für diesen Patienten (s. u.) und die orange Linie stellt den Cut-Off (c-Kriterium nach Jacobson, &

Truax, 1991) zwischen funktionalen und nonfunktionalen OQ-45 Werten dar. Höhere OQ-45 Werte stehen immer für stärkere Beeinträchtigung.

In der linken oberen Box werden patienten- und behandlungsbezogene Daten angezeigt sowie der verwendete Fragebogen und der ausgewählte Algorithmus. Der Algorithmus bestimmt, mit welchem erwarteten Therapieverlauf der Therapieverlauf eines individuellen Patienten X im Feedbackbericht verglichen wird. Für den OQ-45 gibt es zwei Algorithmen: „rational algorithm“ und „empirical algorithm“ (Finch, Lambert, & Schaalje, 2001; Lambert, Whipple, Bishop et al., 2002; Spielmans et al., 2006). Der „rational algorithm“ basiert auf Experteneinschätzungen darüber, welche Veränderung im OQ-45 in welcher Therapiephase in Abhängigkeit vom OQ-45 Eingangswert zu erwarten ist. Der in der hier vorgelegten Arbeit verwendete und in Abbildung 1 dargestellte „empirical algorithm“ basiert auf einer Normstichprobe von 11.492 Patienten, die in verschiedenen psychotherapeutischen Settings in den U.S.A behandelt wurden. Zur Generierung des erwarteten Therapieverlaufs eines individuellen Patienten X werden diejenigen Patienten aus dieser Normstichprobe herangezogen, die einen ähnlich hohen OQ-45 Eingangswert haben wie Patient X. Jeder in Abhängigkeit vom OQ-45 Eingangswert ermittelte erwartete Therapieverlauf beinhaltet mind. 220 Patienten aus der erwähnten Normstichprobe (Finch et al., 2001). Im OQ-Analyst wird nur der OQ-45 Eingangswert zur Bestimmung der erwarteten Therapieverläufe herangezogen, da der OQ-45 Eingangswert sowie die Veränderung von Sitzung 1 bis Sitzung 3 bereits so viel Varianz des finalen Outcomes aufklärten, dass der Einbezug weiterer Variablen (z. B. Therapieverfahren, demographische Variablen) nicht mehr als 1% zusätzliche Varianzaufklärung brachte (Lambert, 2013).

In der rechten oberen Box wird pro Behandlungswoche gezeigt, wie hoch die Belastung des Patienten aktuell und im Vergleich zu Therapiebeginn im OQ-45 ist. Zudem wird pro Behandlungswoche angegeben, ob sich der Patienten seit Therapiebeginn reliabel - gemäß dem „Reliable Change Index“ (RCI; Jacobson, & Truax, 1991) - verändert hat. Außerdem zeigen unter „Alert Status“ - wieder pro Behandlungswoche - verschiedene Farben eine Bewertung des Therapieverlaufs des individuellen Patienten X im Vergleich zu seinen erwarteten Therapieverlauf (hier durch den „empirical algorithm“ bestimmt):

- Ein „Yellow“ Signal erscheint, wenn der Patient um mind. eine Standardabweichung vom erwarteten Therapieverlauf in die negative Richtung abweicht.
- Ein „Red“ Signal wird angezeigt, wenn der Patient so stark vom erwarteten Therapieverlauf in die negative Richtung abweicht, dass nur 10% derjenigen Patienten,

auf denen der erwartete Therapieverlauf basiert, solch einen negativen OQ-45 Wert in der aktuellen Behandlungswoche zeigten.

- Ein „Green“ Signal wird rückgemeldet, wenn der Patient nicht mind. um eine Standardabweichung in die negative Richtung vom erwarteten Therapieverlauf abweicht.
- Ein „White“ Signal wird präsentiert, sobald der Patient einen funktionalen Wert im OQ-45 erreicht und gleichzeitig nicht mind. um eine Standardabweichung negativ vom erwarteten Therapieverlauf abweicht.
- Zusätzlich zum „Green“ und „White“ Signal kann ein „Blue“ Signal erscheinen. Ein „Blue“ Signal bedeutet, dass der Patient so stark vom erwarteten Therapieverlauf in die positive Richtung abweicht, dass in der aktuellen Behandlungswoche solch einen positiven OQ-45 Wert nur 10% derjenigen Patienten erreichten, auf denen der erwartete Therapieverlauf basiert.

In bisherigen Arbeiten wurden Patienten mit mind. einem „Yellow“ oder „Red“ Signal als „Not-On-Track“ (NOT) Patienten und Patienten mit nur „Green“ oder „White“ (bzw. „Blue“) Signalen als „On-Track“ (OT) Patienten bezeichnet (Lambert, 2010). Diese Klassifikation wird in dieser Arbeit beibehalten. Lambert, Whipple, Bishop et al. (2002) und Spielmans et al. (2006) haben untersucht, wie die OT bzw. NOT Klassifikationen des hier verwendeten „empirical algorithm“ während der Therapie mit reliabler Verschlechterung (gemäß RCI nach Jacobson, & Truax, 1991) bzw. nicht reliabler Verschlechterung (reliable Verbesserung sowie nicht reliable Veränderung gemäß RCI) zu Therapieende zusammenhängen. Es zeigte sich, dass 80% - 100% aller Patienten mit reliabler Verschlechterung zu Therapieende während der Therapie als NOT klassifiziert wurden (Hits). Jedoch zeigten bis zu 80% der Patienten, die während der Therapie als NOT klassifiziert werden, keine reliable Verschlechterung zu Therapieende (false alarms). Diese hohe false alarm Rate erscheint jedoch nicht allzu problematisch vor dem Hintergrund, dass eine NOT Klassifikation höchstens dazu führen kann, dass der Therapeut sein therapeutischen Vorgehen reflektiert und sich daraus keine gefährlichen Konsequenzen für den Patienten ergeben wie es z. B. in manchen Bereichen der Medizin der Fall sein könnte, wenn aufgrund eines Warnsignals ein invasiver Eingriff eingeleitet werden würde.

Geeignet ist der „empirical algorithm“ den berichteten Ergebnissen zufolge dafür, Patienten, die sich zu Therapieende reliabel verschlechtern, frühzeitig schon während der Behandlung zu identifizieren. Die Hit Rate des „empirical algorithm“ von 80% - 100% übersteigt die von Therapeuten, die - wie mehrere Studien gezeigt haben - mit der Vorhersage von Verschlechterungen ihre Schwierigkeiten haben (Chapman et al., 2012; Hannan et al., 2009;

Hatfield, McCullough, Frantz, & Krieger, 2010): Von 26 Patienten, die sich zu Behandlungsende reliabel verschlechterten, konnten Therapeuten dies in der Studie von Hannan et al. (2005) nur bei 1 Patienten frühzeitig schon während der Behandlung vorhersagen (Hit Rate = 3.8%). Diese Befunde stehen in Einklang mit dem Forschungsergebnis, dass statistische Einschätzungen im „mental health“ Bereich höhere Genauigkeit bieten als klinische Einschätzungen (Ægisdóttir et al., 2006).

Im Feedbackbericht zum Therapieverlauf (Abbildung 1) werden auch noch die aktuellen Antworten zu den fünf von den Entwicklern des OQ-Analyst als besonders wichtig erachteten Items des OQ-45 angezeigt sowie Vergleiche der aktuellen OQ-45 Skalenwerte mit Normstichproben. Außerdem gibt es je nach „Alert Status“ eine entsprechende „Feedback Message“ im Sinne einer Behandlungsempfehlung am Ende des Feedbackberichts.

Neben Feedback zum Therapieverlauf bietet der OQ-Analyst auch CST. Die CST im OQ-Analyst bestehen aus zwei Komponenten: Erstens erhält der Therapeut Feedback über die Ergebnisse seines Patienten im Fragebogen ASC, damit der Therapeut die im ASC erhobenen Faktoren (Therapiebeziehung, soziale Unterstützung, Motivation, Lebensereignisse) als mögliche Einflussfaktoren auf den Therapieverlauf seines Patienten in Erwägung zieht. Zweitens soll der Therapeut durch einen „Decision Tree“ angeleitet werden, mit problematisch ausgeprägten ASC Ergebnissen therapeutisch adäquat umzugehen. Im Folgenden wird zuerst die eine CST Komponente (ASC Feedback) und danach die andere CST Komponente („Decision Tree“) vorgestellt.

Sobald ein ASC in den OQ-Analyst übertragen wurde, kann ein Feedbackbericht zu den ASC Ergebnissen erstellt werden. Abbildung 2 zeigt beispielhaft einen solchen Feedbackbericht zu den ASC Ergebnissen.

Name: ID: Session Date: Clinician: Diagnosis: Unknown Diagnosis Instrument: ASC Display Interventions Handout	Subscales Therapeutic Alliance: 37 RED Social Support: 29 RED Motivation: 32 RED Life Events: 29
Therapeutic Alliance: RED It is advisable that you address your relationship with the client. Please click 'Display Interventions Handout' button for more information. 1. I felt cared for and respected as a person. Neutral 2. I felt my therapist understood me. Neutral 3. I thought the suggestions my therapist made were useful. Slightly Agree 4. I felt like I could trust my therapist completely. Neutral 5. I was willing to share my innermost thoughts with my therapist. Neutral 6. I felt there was a breakdown in the relationship with my therapist. Neutral 7. I felt like my therapist disapproved of me. Slightly Disagree 9. My therapist seemed to be glad to see me. Neutral 10. My therapist and I seemed to work well together to accomplish what I want. Slightly Agree 11. My therapist and I had a similar understanding of my problems. Neutral	Social Support: RED It is advisable that you address the client's support outside of therapy. Please click 'Display Interventions Handout' button for more information. 16. There was a special person who was around when I was in need. Neutral 17. There was a special person with whom I could share my joys and sorrows. Neutral 18. I could get material support if needed (like: money, food, transportation, child care, tools, repairs, health care, legal advice, etc.) Strongly Disagree 19. I had support from social groups (like: church, school, AA, clubs, etc.) Strongly Disagree 20. I felt accepted by someone other than my therapist. Neutral
Motivation: RED It is advisable that you address your client's motivation in therapy. Please click 'Display Interventions Handout' button for more information. 28. I have no desire to work out my problems. Neutral 30. Through therapy I am taking more responsibility for changing my life. Neutral 31. I am in therapy because someone is requiring it of me. Neutral	Life Events:

Abbildung 2: Feedbackbericht zu den Ergebnissen im „Assessment for Signal Cases“ (ASC).
Mit freundlicher Genehmigung von OQ-Measures.

Im Feedbackbericht zu den ASC Ergebnissen (Abbildung 2) werden links oben patienten- und behandlungsbezogene Daten und rechts oben die Skalenwerte der aktuellen ASC Messung rückgemeldet. Höhere Skalenwerte stehen im ASC für positivere Ausprägungen der erhobenen Faktoren. Im Gegensatz zum Feedback über den Therapieverlauf (OQ-45, Abbildung 1) wird zu den ASC Ergebnissen im OQ-Analyst kein Verlauf rückgemeldet, sondern pro Behandlungswoche die aktuellen Ausprägungen der ASC Skalen. Ein „RED“ Signal erscheint bei den ASC Skalen, wenn der Skalenwert eines Patienten unter einen Cut-Off Wert fällt (Lambert et al., 2007), der einem Prozentrang kleiner gleich 20% entspricht, d. h. 80% einer Normstichprobe zeigten höhere (bedeutet im ASC bessere) Scores auf dieser Skala (Simon et al., 2013). Ein „RED“ Signal in einer ASC Skala soll dem Therapeuten anzeigen, dass dieser

ASC Faktor einen negativen Einfluss auf den Therapieverlauf des Patienten haben könnte. Zudem werden diejenigen ASC Items rückgemeldet, in denen der Patient unter einem Cut-Off Wert (Prozentrang kleiner gleich 20%) scort. Dies soll den Therapeuten darüber informieren, welche konkreten Probleme der Patient im ASC angibt und was somit therapeutische Ansatzpunkte sein könnten. Unter „Display Interventions Handout“ kann der „Decision Tree“ - die zweite CST Komponente im OQ-Analyst - aufgerufen werden, der den Therapeuten beim therapeutisch adäquaten Umgang mit problematisch ausgeprägten ASC Ergebnissen (im Sinne eines „RED“ Signals) unterstützen soll. In diesem „Decision Tree“ wird der Therapeut angeleitet, zuerst die ASC Ergebnisse zur Therapiebeziehung zu prüfen. Sollte sich dort ein problematisches Ergebnis zeigen, so wird der Therapeut auf evidenzbasierte Interventionen zur Verbesserung der Therapiebeziehung verwiesen. Im nächsten Schritt soll der Therapeut die Motivation des Patienten prüfen. Bei problematischem Ergebnis in der ASC Motivationskala, kann sich der Therapeut Informationen zu evidenzbasierten Interventionen zur Motivationssteigerung anzeigen lassen. Im dritten Schritt erfolgt das gleiche Vorgehen für die soziale Unterstützung und im vierten Schritt sollen die Lebensereignisse genauer betrachtet werden, wobei es laut OQ-Analyst noch keine evidenzbasierten Interventionen zur Reduktion von kritischen Lebensereignissen gibt. Daher werden dem Therapeuten keine Interventionen bzgl. Lebensereignissen vorgeschlagen¹. In den letzten beiden Schritten des „Decision Tree“ wird der Therapeut zu Überlegungen bzgl. des Behandlungsplans und der Medikation angeregt. Weiterführende Informationen zum verwendeten „Decision Tree“ sind zu finden bei Lambert et al. (2007) sowie bei Harmon, Hawkins, Lambert, Slade, und Whipple (2005) zu einer früheren Version des „Decision Tree“.

Die CST des OQ-Analyst wurden für NOT Patienten entwickelt und bei der Nutzung des OQ-Analyst in der klinischen Routineversorgung ist vorgesehen, den ASC und den „Decision Tree“ nur bei Patienten zu verwenden, wenn sie als NOT klassifiziert werden. Diesem Vorgehen liegt die Idee zugrunde, dass bei Patienten mit deutlich negativer Abweichung vom erwarteten Therapieverlauf eine ausführlichere Problemfallanalyse notwendig ist. Im Rahmen dieser Problemfallanalyse sollen zuerst mögliche Einflussfaktoren auf den negativen Therapieverlauf identifiziert werden (mit dem ASC) und danach soll der Therapeut angeleitet werden, mit problematisch ausgeprägten ASC Ergebnissen konstruktiv umzugehen (anhand des „Decision Tree“). Die vier Bereiche, die der ASC erfasst (Therapiebeziehung, soziale Unter-

¹ Zur Förderung des therapeutisch adäquaten Umgangs mit kritischen Lebensereignissen könnten Therapeuten hier jedoch auf evidenzbasierte Copingstrategien (z. B. Garnefski, Kraaij, & Spinhoven, 2001) hingewiesen werden, die sie Patienten mit kritischen Lebensereignissen in der Therapie vermitteln könnten.

stützung, Motivation, Lebensereignisse), wurden als mögliche Einflussfaktoren auf den Therapieverlauf ausgewählt, da sie in empirischen Studien mit dem Therapieverlauf bzw. dem Prä-Post Outcome verbunden waren (Bankoff, & Howard, 1992; Horvath, Del Re, Flückiger, & Symonds, 2011; Pelletier, Tuson, & Haddad, 1997; Sexton, 1996).

Zusammengefasst integriert der OQ-Analyst die drei Elemente, die auch im Rahmen der EBM gefordert werden: 1) Verfügbare wissenschaftliche Evidenz (Evidenzbasierte Interventionen zum Umgang mit problematischen ASC Ergebnissen im „Decision Tree“), 2) die Patientenperspektive (Feedback über den Therapieverlauf und über mögliche Einflussfaktoren aus Patientensicht) und 3) die klinische Expertise des Therapeuten (Entscheidung über Integration des Feedbacks und des „Decision Tree“ in die Behandlung bleibt dem Therapeuten überlassen).

Der OQ-Analyst ist das bisher am meisten erforschte Monitoring- und Feedbacksystem (Castonguay et al., 2013). Die bisherigen Studien, in denen die Algorithmen des OQ-45 zur Klassifikation von NOT bzw. OT Patienten eingesetzt wurden und in denen Therapeuten für NOT und OT Patienten regelmäßig Feedback über den Therapieverlauf bekamen sowie zusätzlich für NOT Patienten auch CST erhielten, fanden statt in Beratungsstellen (Harmon et al., 2007; Slade, Lambert, Harmon, Smart, & Bailey, 2008; Whipple et al., 2003), ambulanten Psychotherapien (Simon, Lambert, Harris, Busath, & Vazquez, 2012), ambulanten Behandlungen von Drogenmissbrauch (Crits-Christoph et al., 2012) und stationären Behandlungen von Essstörungen (Simon et al., 2013). Zusätzlich wurden Studien durchgeführt, die den Einfluss von Feedback über den Therapieverlauf im OQ-45 untersuchten, ohne den Therapeuten CST für NOT Patienten zur Verfügung zu stellen (Amble, Gude, Stubdal, Andersen, & Wampold, 2014; Hansson, Rundberg, Österling, Öjehagen, & Berglund, 2013; Lambert et al., 2001; Lambert, Whipple, Vermeersch et al., 2002). Überdies wurden zwei Meta-Analysen über die bisherigen Ergebnisse der Primärstudien veröffentlicht (Lambert et al., 2003; Shimokawa, Lambert, & Smart, 2010). Die Ergebnisse dieser Studien, die für das hier berichtete Projekt in stationärer Psychosomatik von besonderer Relevanz sind, werden in den Publikationen unter III. detaillierter berichtet. Kurz zusammengefasst, zeigten diese Studien folgende Feedbackeffekte auf den Patientenoutcome: Erstens, Feedback über den Therapieverlauf (ohne CST) an den zuständigen Behandler verbesserte den Prä-Post Outcome von NOT Patienten gegenüber „Treatment-As-Usual“ (TAU) ohne Feedback (Lambert et al., 2001; Lambert, Whipple, Vermeersch et al., 2002). Zweitens, die Kombination aus Feedback über den Therapieverlauf

und CST an den zuständigen Behandler verbesserte den Prä-Post Outcome von NOT Patienten sowohl gegenüber Feedback über den Therapieverlauf ohne CST (Harmon et al., 2007; Slade et al., 2008; Whipple et al., 2003) als auch gegenüber TAU (Crits-Christoph et al., 2012; Harmon et al., 2007; Simon et al., 2012; Slade et al., 2008; Whipple et al., 2003). Drittens, der Prä-Post Outcome von OT Patienten konnte durch Feedback über den Therapieverlauf (ohne CST) in drei Primärstudien gegenüber TAU nicht gesteigert werden (Lambert et al., 2001; Lambert, Whipple, Vermeersch et al., 2002; Whipple et al., 2003), in einer jedoch schon (Harmon et al., 2007). Viertens, wenn Feedback über den Therapieverlauf (ohne CST) nicht nur den zuständigen Behandler, sondern auch dem entsprechenden Patienten zur Verfügung gestellt wurde, besserte dies den Prä-Post Outcome gegenüber Feedback über den Therapieverlauf nur an den zuständigen Behandler zwar in einer Primärstudie (Hawkins et al., 2004), allerdings nicht in zwei Folgestudien (Harmon et al., 2007; Slade et al., 2008).

II. 3. ICD-10-Symptom-Rating

Neben der im vorherigen Gliederungspunkt beschriebenen Monitoring- und Feedbacksoftware OQ-Analyst, mit dessen Hilfe in diesem Projekt Feedbackberichte für die Therapeuten zu den Monitoringdaten des OQ-45 und ASC erstellt wurden, wurde in der vorliegenden Arbeit im Rahmen des ROM auch das ICD-10 Symptom-Rating (ISR; Tritt et al., 2008; in Vorbereitung) eingesetzt. Das ISR ist ein Selbsteinschätzungsfragebogen bestehend aus 29 Items, die in enger Anlehnung an diagnostische Kriterien nach Kapitel V des ICD-10 formuliert wurden. Das ISR umfasst eine Gesamtskala sowie fünf Syndromskalen (Depressionsskala, Angstskala, Zwangsskala, Somatisierungsskala, Essstörungsskala). Darüber hinaus liefern die 12 Zusatzitems des ISR noch erste Hinweise auf das Vorliegen einer Reihe weiterer psychischer Syndrome. Das ISR hat als Zielsetzung, die psychische Symptomatik für Status- und Veränderungsmessungen zu evaluieren. Im ersten Entwicklungsschritt wurde das ISR hinsichtlich verschiedener testpsychologischer Gütekriterien untersucht (z. B. Fischer, Schirmer, Tritt, Klapp, & Fliege, 2011; Fischer, Tritt, Klapp, & Fliege, 2010; Tritt et al., 2010). Im zweiten Entwicklungsschritt, an dessen Umsetzung derzeit gearbeitet wird, soll das ISR auf der Basis computergestützten adaptiven Testens eine Screening-Funktion zur Verbesserung der ICD-10-Diagnostik übernehmen. Hierzu ist derzeit angedacht, dass die Diagnostiker im Gespräch mit den Patienten ein zweistufiges Expertenrating durchführen:

Die erste Stufe des Expertenratings basiert auf den ISR Werten eines Patienten zur Eingangsmessung: Der Diagnostiker erhält Feedback darüber, bei welchen der im ISR erfass-

ten Syndromen aufgrund der ISR Eingangsmessung ein Verdacht auf das Vorliegen einer psychischen Störung besteht. Für die aufgrund der ISR Werte im Verdacht stehenden psychischen Störungen erhält der Diagnostiker dann eine Auflistung der Diagnosekriterien (nach Kapitel V des ICD-10) rückgemeldet, deren Vorhandensein der Diagnostiker beim Patienten prüfen und die zutreffenden Diagnosekriterien bewerten soll.

In der zweiten Stufe des Expertenratings erhält der Diagnostiker eine Liste aller bislang noch nicht gesichteten psychischen Syndrome, die im ICD-10 erfasst werden, und gibt an, bei welchen psychischen Syndromen aus seiner Sicht ein Verdacht auf das Vorliegen einer psychischen Störung beim Patienten besteht. Bei den ausgewählten Störungen wird der Diagnostiker anschließend durch die diagnostischen Algorithmen nach Kapitel V des ICD-10 gelotst. Zum Abschluss soll der Diagnostiker Feedback über die sich aus dem Expertenrating ergebenden psychischen Diagnosen nach Kapitel V des ICD-10 erhalten, die Ausschlusskriterien nach ICD-10 prüfen, die Diagnosen bestätigen und die Haupt- sowie die Nebendiagnosen definieren.

II. 4. Forschungsagenda

Die in den vorherigen Kapiteln näher beschriebenen Fragebögen OQ-45, ASC und ISR sowie die Monitoring- und Feedbacksoftware OQ-Analyst wurden in dieser Arbeit wie folgt in zwei psychosomatischen Kliniken eingesetzt: Die Patienten wurden bei Aufnahme entweder in die experimentelle Feedbackgruppe randomisiert oder in die TAU Kontrollgruppe. Monitoring wurde für die Patienten beider Gruppen wöchentlich (jeden Freitag) mit den beiden Fragebögen des OQ-Analyst (OQ-45 und ASC) und dem ISR realisiert. Die wöchentlich erhobenen OQ-45 und ASC Daten wurden in den OQ-Analyst übertragen. Danach wurden für Patienten der experimentellen Feedbackgruppe die Feedbackberichte des OQ-Analyst ausgedruckt und dem für den jeweiligen Patienten zuständigen Einzeltherapeuten ausgehändigt. Dabei wurden sowohl die Feedbackberichte zum Therapieverlauf (OQ-45, Abbildung 1) als auch die Feedbackberichte zu den möglichen Einflussfaktoren auf den Therapieverlauf (ASC, Abbildung 2) rückgemeldet. Für ihre Patienten der TAU Kontrollgruppe bekamen die Therapeuten keine Feedbackberichte des OQ-Analyst. Den unter II.2. beschriebenen „Decision Tree“ des OQ-Analyst erhielten die Therapeuten zu Beginn des Projekts, als ihnen die Interpretation der Feedbackberichte transparent gemacht wurde. Den Therapeuten wurde bei dieser Gelegenheit auch die Telefonnummer des Projektleiters (Autor dieser Arbeit) ausgehändigt mit dem Hinweis, sich bei Fragen zu den Feedbackberichten oder zum Gesamtprojekt telefonisch melden

zu können. Zum wöchentlichen Therapieverlauf im ISR wurden den Therapeuten in diesem Projekt keine Feedbackberichte zur Verfügung gestellt, da bisher noch keine Software existiert, die die ISR Verläufe individueller Patienten mit erwarteten Therapieverläufen vergleicht. Jedoch hatten die Therapeuten Zugriff auf Feedbackberichte zu den ISR Statusmessungen bei Aufnahme und Entlassung für all ihre Patienten (EG und CG), da solche Feedbackberichte im Rahmen der routinemäßig durchgeführten Qualitätssicherung in den Kliniken erstellt werden.

Im Folgenden werden die Fragestellungen der vier in diese Dissertation eingegangenen Publikationen skizziert, bevor die dazugehörigen Publikationen im Kapitel III. präsentiert werden.

Fragestellung 1 / Publikation 1: In Beratungsstellen (Harmon et al., 2007; Slade et al., 2008; Whipple et al., 2003), ambulanten Behandlungen von Drogenmissbrauch (Crits-Christoph et al., 2012) und ambulanten Psychotherapien (Simon et al., 2013) konnte der Outcome von NOT Patienten bei Behandlungsende gegenüber TAU verbessert werden, wenn Therapeuten Feedback über den Therapieverlauf und CST für ihre NOT Patienten erhielten. In der ersten Publikation (Probst et al., 2013) wurde der Fragestellung nachgegangen, ob sich dieses Ergebnis auch in der stationären Psychosomatik replizieren lässt.

Fragestellung 2 / Publikation 2: In früheren Studien zeigten sich widersprüchliche Ergebnisse, zu der Fragestellung ob TAU auch für OT Patienten verbessert werden kann, wenn Therapeuten Feedback über den Therapieverlauf (*ohne CST*) für ihre OT Patienten erhalten (Harmon et al., 2007; Lambert et al., 2001; Lambert, Whipple, Vermeersch et al., 2002; Shimokawa et al., 2010; Whipple et al., 2003). Daher wurde in der zweiten Publikation (Probst, Lambert, Dahlbender, Loew, & Tritt, 2014) die Fragestellung untersucht, ob der Therapieprozess von OT Patienten gegenüber TAU in der stationären Psychosomatik optimiert werden kann, wenn Therapeuten die Kombination aus Feedback über den Therapieverlauf *plus* CST für ihre OT Patienten erhalten.

Fragestellung 3 / Publikation 3: In den meisten Studien zur Evaluation der stationären Psychosomatik wurde das Prä-Post Behandlungsergebnis untersucht (Liebherz, & Rabung, 2013; Steffanowski, Löschmann, Schmidt, Wittmann, & Nübling, 2007). In einigen Studien wurden auch Katamnesen (z. B. Franke, Hoffmann, & Frommer, 2005; Probst et al., 2009) oder eine Zwischenerhebung während der Behandlung (z. B. vier Wochen nach Aufnahme) durchge-

führt (z. B. Bauer, 2004; Franke et al., 2005; Haase et al., 2008). Die wenigen Studien zur stationären Psychosomatik, die behandlungsrelevante Kriterien im wöchentlichen Verlauf erhoben haben (Berking et al., 2006; Dinger & Schauenburg, 2010; Melchior, 2011), berichteten nicht, ob sich verschiedene behandlungsrelevante Kriterien differentiell verändern. Deshalb wurde im Rahmen der Fragestellung der dritten Publikation (Probst, Loew, Dahlbender, & Tritt, 2014) untersucht, welche behandlungsrelevanten Kriterien - erfasst durch den OQ-45 und das ISR - sich im Verlauf der ersten vier Behandlungswochen in welchem Ausmaß verändern und wie hoch die Eingangsbelastungen in diesen Kriterien sind.

Fragestellung 4 / Publikationen 4: Bei Verwendung des OQ-Analyst in der klinischen Routinerversorgung ist vorgesehen, den ASC nur bei Patienten mit NOT Klassifikation zu verwenden, um mögliche Einflussfaktoren auf die NOT Klassifikation zu eruieren. Da die vier im ASC erhobenen Faktoren in früheren Studien mit dem Prä-Post Outcome bzw. dem Therapieprozess zusammenhingen, wurde bei der Entwicklung des OQ-Analyst davon ausgegangen, dass sie auch mit bedeutsamen negativen Verschlechterungen (d. h. NOT Klassifikationen) assoziiert sein könnten. Diese Annahme wurde bei Probst et al. (under re-review) empirisch untersucht. Außerdem wurden in dieser vierten Publikation Zusammenhänge zwischen bedeutsamen positiven Verbesserungen („Blue“ Signale im OQ-Analyst) und den vier im ASC erhobenen Faktoren analysiert.

III. Publikationen

III. 1. Publikation 1

Probst, T., Lambert, M. J., Loew, T. H., Dahlbender, R. W., Gollner, R., & Tritt, K. (2013). Feedback on patient progress and clinical support tools for therapists: Improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice. *Journal of Psychosomatic Research*, 75, 255-261.

doi: 10.1016/j.jpsychores.2013.07.003

Link: <http://www.sciencedirect.com/science/article/pii/S0022399913002699>



Feedback on patient progress and clinical support tools for therapists: Improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice[☆]

Thomas Probst^{a,b}, Michael J. Lambert^{c,*}, Thomas H. Loew^a, Reiner W. Dahlbender^d, Richard Göllner^e, Karin Tritt^a

^a University Clinic of Regensburg, Department of Psychosomatics, Germany

^b University of Regensburg, Clinical Psychology and Psychotherapy, Germany

^c Brigham Young University, Department of Psychology, United States

^d University Clinic of Ulm, Clinic for Psychosomatic Medicine and Psychotherapy, Germany

^e Eberhard Karls University of Tübingen, Center for Educational Science and Psychology, Germany

ARTICLE INFO

Article history:

Received 6 February 2013

Received in revised form 2 July 2013

Accepted 6 July 2013

Keywords:

Inpatient treatment

Outcome

Psychosomatics

Psychotherapy

Routine clinical care

Treatment failure

ABSTRACT

Objectives: Although psychosomatic in-patient treatment is effective, 5–10% of the patients deteriorate. Providing patient progress feedback and clinical support tools to therapists improves the outcome for patients at risk of deterioration in counseling, outpatient psychotherapy, and substance abuse treatment. This study investigated the effects of feedback on psychosomatically treated in-patients at risk of treatment failure.

Methods: At intake, all patients of two psychosomatic clinics were randomized either into the experimental group or the treatment-as-usual control group. Both groups were tracked weekly with the "Outcome Questionnaire" (OQ-45) measuring patient progress and with the clinical support tool "Assessment of Signal Cases" (ASC). Therapists received feedback from both instruments for all their experimental group patients. "Patients at risk" were defined as patients who deviated from expected recovery curves by at least one standard deviation. Of 252 patients, 43 patients were at risk: 23 belonged to the experimental group, 20 to the control group. The feedback effect was analyzed using a level-2-model for discontinuous change, effect size (d), reliable change index (RCI), and odds ratio for reliable deterioration.

Results: For patients at risk, the experimental group showed an improved outcome on the OQ-45 total scale compared to the control group ($p < 0.05$, $d = 0.54$). By providing feedback, the rate of reliably deteriorated patients at risk was reduced from 25.0% (control group) to 8.7% (experimental group) – odds ratio = 0.29. All reliably improved patients at risk belonged to the experimental group.

Conclusion: Feedback improves the outcome of patients at risk undergoing psychosomatic in-patient treatment.

© 2013 Elsevier Inc. All rights reserved.

Introduction

In psychosomatic in-patient therapy, a multi-professional team uses multi-modal and multi-method treatments on patients with mental and behavioral disorders too severe for outpatient treatment [1]. Many studies have proven the effectiveness of psychosomatic in-patient treatment (e.g. [2–4]). Nevertheless, 5–10% of the patients with mental health problems deteriorate during psychosomatic in-patient treatment [5,6]. Inspired by Bergin's first study on deterioration effects in psychotherapy [7], research has identified patient-, therapist-, and treatment-related variables that might put patients

at risk of deterioration [8,9]. Managing these risk factors and reducing deterioration rates are relevant issues for clinical practice and for psychotherapy health service research [10]. To prevent deterioration, risk of treatment failure must be detected before patients at risk withdraw from treatment, so that countermeasures can be implemented. To detect risk of deterioration early on in psychotherapy, algorithms (rational and empirical algorithms) were formulated [11–13] using the self-report "Outcome Questionnaire" (OQ-45) [14]. Based on the initial severity of psychological distress (OQ-45 total score) of a given patient, these algorithms provide an expected individual recovery curve for this patient. For every further measurement point, the individual patient's progress is then tracked against the expected recovery curve. If a patient's score deviates negatively from the expected recovery curve as defined by the algorithms, the patient is classified as at risk of treatment failure. While patients not labeled as at-risk patients have deterioration rates between 0.3% and 1.3%, at-risk patients show deterioration rates between 8.4% and 19.4%

[☆] The work was conducted at the Psychosomatics Department of the Hospital in Donaustauf, Germany, and the Psychosomatic Hospital "Am Schönen Moos" in Bad Saulgau, Germany.

* Corresponding author at: Brigham Young University, Department of Psychology, 272 TLRB, Provo, 84602, USA. Tel.: +1 801 422 6480.

E-mail address: Michael_Lambert@byu.edu (M.J. Lambert).

(results for the empirical algorithm [12,13]). Hence, patients classified as at risk have a much higher risk of leaving psychotherapy deteriorated. Furthermore, using the OQ-45 algorithms to identify patients at risk proves to be a much more accurate predictor of negative treatment outcome than subjective evaluations provided by clinicians [15]. The software OQ-Analyst [16], rated as an evidence-based practice with excellent reviews for its training material and easy implementation [17], can be employed to classify patients at risk by comparing individual progress against expected curves. To do so, the patient's progress should be tracked on a weekly basis using the OQ-45. Using the OQ-Analyst, therapists receive weekly feedback on patient progress, the expected recovery curve, and a warning signal if the patient is labeled as at risk. Tracking the patient progress, also known as monitoring, and providing (patient progress) feedback, originate from patient-focused research, aiming at improving the outcome of individual psychotherapies [18].

Providing patient progress feedback to therapists has been shown to improve the outcome of patients at risk in counseling [19,20], outpatient psychotherapy [21], and psychiatric care [22]. At-risk patients whose therapists received patient progress feedback also stayed longer in treatment [23], probably contributing to the improved outcome. The outcome of patients at risk can be further improved when therapists not only receive feedback on patient progress, but also are given clinical support tools (CST). CST can be described as an empirically based problem-solving strategy, helping therapists to focus on important factors which can influence the outcome of psychotherapy [24]. The CST instrument within the OQ-Analyst is the so-called "Assessment of Signal Cases" (ASC) and includes feedback on therapeutic alliance, motivation for change, social support, critical life events, and a decision tree for handling related problems [25]. The existing studies involving feedback on patient progress and CST took place in counseling [24,26,27], outpatient psychotherapy [28], substance abuse treatment [29], and in-patient eating-disorder therapy [30]. This study employed the OQ-Analyst for the first time in German psychosomatic in-patient clinics to investigate the effects of feedback on patient progress and CST for psychosomatically treated in-patients at risk of a negative outcome.

Methods

Questionnaires

Outcome Questionnaire (OQ-45): The OQ-45 was used to track patient progress. This self-report questionnaire includes 45 items measured on a five-point Likert scale. Each item pertains to the previous week. There are three subscales (symptom distress, interpersonal problems, and social role performance) and one total scale; the higher the values, the higher the distress level. This study used the total scale which comprises all 45 items and indicates the general severity of psychological distress. The total scale of the German OQ-45 version has an internal consistency of 0.93 and a retest-reliability of $r = 0.88$ [31].

Assessment of Signal Cases (ASC): The ASC served as the CST instrument. This self-report questionnaire consists of 40 items. Each item is answered on a five-point Likert scale and covers the previous week. The ASC consists of four scales: therapeutic alliance, social support, motivation for change, and life events. These four areas are considered to significantly influence patient progress [25]. The ASC used in this study was translated into German via a back-translation method. To ensure the quality of the translation, cognitive debriefing interviews with ten in-patients were done with the German version of the ASC before the study began [32]. In contrast to former studies [24,26–30], the CST was used not only for at-risk patients in this study; wanting to gather further experience and collect data with the instrument, all patients were encouraged to complete the ASC every week during the course of their psychosomatic treatment. Testing the reliability of the four ASC scales of the German ASC version with the internal consistency coefficient Cronbach's alpha, the following coefficients were found for

all patients (not only at-risk patients) with an intake ASC: therapeutic alliance scale: 0.89; social support scale: 0.76; motivation for change scale: 0.78; and life events scale: 0.71.

Study design

All patients of the Psychosomatics Department of the Hospital in Donaustauf, Germany, starting treatment from 10/01/2010 to 07/06/2012 were asked to participate in the study. At the Psychosomatic Hospital "Am Schönen Moos" in Bad Saulgau, Germany, all patients with private health care who entered treatment between 01/16/2012 to 04/06/2012 were encouraged to take part in the study. Both clinics provide multi-modal and multi-method psychosomatic in-patient treatment with more than 24 treatment hours per week, including individual and group psychotherapy, relaxation and mindfulness training, physical activity therapy, creative therapy (art, dance, music), and – if required – crisis intervention, visitations by nurses, medical consultations, etc. The multi-professional teams consisted mainly of psychologists, physicians, and nurses. At clinic intake, all patients (not only at-risk patients) were randomized either into the experimental group or the treatment-as-usual control group. All patients of both treatment groups were monitored: Every patient was asked to complete the OQ-45 and the ASC (paper and pencil versions) each week sometime between Friday and Sunday during the course of their stay. After the OQ-45 and ASC data was transferred to the OQ-Analyst, the feedback reports of the OQ-Analyst were printed on the following Monday. The feedback reports were given on a weekly basis to the therapists in a closed envelope only for the patients of the experimental group. The therapists received feedback only for those experimental group patients that they worked with in individual therapy. The feedback on patient progress and CST was provided to therapists for all experimental group patients, not only for at-risk patients. As feedback on patient progress requires at least two OQ-45 scores, the therapists received the weekly feedback starting with the second measurement point of a given experimental group patient. The therapists could freely choose to discuss the feedback information with the patient, the clinic team and/or supervisors. Before the study commenced, interpretation of the feedback reports was explained in detail to the therapists. They were also motivated to view the feedback information not as prescriptive, but as a possibility to broaden clinical problem solving with their patients.

Focusing on the effects of feedback on patients at risk of deterioration, the empirical algorithm of the OQ-Analyst was applied to classify patients at risk: If the patient's progress deteriorated by at least one standard deviation from his expected recovery curve at any week during treatment, the patient was considered as being at risk and the therapist received a warning signal in the weekly feedback report, implying that this patient is at risk of treatment failure. The expected recovery curve of a given patient is based upon a norm derived from a large group of U.S. patients with a similar intake OQ-45 total score, calculated with at least 220 patients for each specific expected recovery curve [11]. Patients without an intake OQ-45 score and patients without at least three measurement points were excluded from this study: The intake OQ-45 score constitutes the baseline of the expected recovery curve. Since giving feedback requires at least two measurement points, three or more measurement points are a prerequisite to study the effects of feedback on patient progress.

Statistical analysis

The statistical analysis was performed with SPSS 19.0 FP2, applying a significance level of ≤ 0.05 . All tests were performed two-tailed. Means (M) and standard deviations (SD) were calculated for the sample description. To explore differences between the two treatment groups in regard to age, gender, education, comorbidity, treatment duration, and assessment points with first at-risk signal,

t-tests were performed for metric variables and Fisher's exact tests (FET) for nominal variables.

For the statistical analysis of the feedback effects on patients at risk, three measurement points were chosen:

- t1 The first measurement point (t1) was the OQ-45 total score from the clinic intake week.
- t2 The OQ-45 total score of the week in which the patient deviated negatively by at least one standard deviation from the expected recovery curve ("patient at risk") was defined as the signal measurement point (t2). For the experimental group, the patient's therapist received a warning signal in the feedback report at t2 (red or yellow alarms [11–13]) implying that the patient is at risk of treatment failure.
- t3 The last measurement point (t3) was the last available OQ-45 total score. As in prior studies [26,27], the last observation carried forward (LOCF) method was applied as a conservative estimation of treatment outcome. The change occurring from t1 to t3 was used to assess outcome.

The change pattern from t1 to t2 was expected to be similar for both treatment groups, but – with a warning signal given at t2 – an improved outcome from t1 to t3 was anticipated for the experimental group. To test these hypotheses, a multi-level model for discontinuous change [33, chapter 6; 34] with maximum likelihood (ML) estimation was used. A multi-level model was chosen instead of the traditional repeated-measures analysis of variance, since its assumptions are less restrictive, e.g. equidistant measurement points are not required [33,34]. Since previous studies have shown that patients at risk feel worse during the time between intake and the signal measurement point (t2), but improve from the signal to the last measurement point (e.g. [21,24,26,27]), a model for discontinuous change seems best to fit patients at risk. Our multi-level model is a level-2-model: The measurement points constitute the first level and the individuals the second level of analysis, whereby differences between experimental and control group represent models at the second level. As described by Gollner and colleagues [34, model 4] for the level-2-model for discontinuous change, the three measurement points (t1, t2, t3) were coded into two contrast variables with the first measurement point being the reference. Simple contrast tests compared the experimental with the control group from t1 to t2 as well as from t1 to t3. The mixed command in SPSS 19.0 FP2 was used for the level-2-model. The repeated subcommand was selected to produce an unstructured variance-covariance matrix.

To assess the effect size (d), Hedges' g [35] was computed for the individual measurement points t1 and t3. The adjusted feedback effect occurring between t1 and t3 was measured by adding up the specific effect sizes resulting for t1 and t3 [35].

To explore the change of each single patient at risk, the reliable change index (RCI) [36] was applied from t1 to t3. According to Lambert and colleagues [31], a patient was classified as "reliably improved" in the total scale of the German OQ-45 if he improved by at least 21 points. He was considered "reliably deteriorated" if he deteriorated 21 or more points and "unchanged" if a change of less than 21 points was found. To further evaluate the occurrence of the event "reliable deterioration" in both groups, the odds ratio for "reliable deterioration" was calculated.

Sample

252 patients with an OQ-45 intake score and at least two further OQ-45 assessments were available for analysis (184 patients without intake score or without at least three measurement points were excluded). A total of 17 therapists, who worked with the 252 patients in individual therapy, participated in the study. 43 of the 252 patients were classified as patients at risk of deterioration (17.1% of 252); 23 of the patients at risk belonged to the experimental group and 20 were participants of the control group. 13 therapists worked with the 43

at-risk patients in individual therapy. Accordingly 4 therapists had no at-risk patients. In one clinic 6.7% of the patients were classified as at risk of treatment failure, while in the other clinic 17.7% of the patients were labeled as at risk. The distribution of at-risk and not-at-risk patients did not differ significantly between the clinics ($p = 0.48$, FET). More data however is not presented for each separate clinic/therapist to protect the confidentiality agreement with the clinics and therapists. The diagnoses of the patients were made by the clinic teams according to chapter F (mental and behavioral disorders) of the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10) [37]. Considering all comorbid F-Diagnoses the patients at risk received, the most frequent diagnoses included depressive disorders F32 and F33 (76.7% of all patients at risk, 78.3% of the experimental group, and 75.0% of the control group), somatoform disorders F45 (58.1% of all patients at risk, 52.2% of the experimental group, and 65.0% of the control group), and anxiety disorders F40 and F41 (20.9% of all patients at risk, 30.4% of the experimental group, and 10.0% of the control group).

Age, gender, education, comorbidity, duration of treatment weeks, and assessment points with first at-risk signal are shown in Table 1. No significant differences between the two treatment groups were found, although the experimental group seemed to be more educated, to have less comorbid disorders, and to stay longer in treatment.

Results

The level-2-model for discontinuous change involved six fixed parameters and six parameters for the unstructured variance-covariance matrix. The six parameters of the unstructured variance-covariance matrix were: variance t1 = 365.66, variance t2 = 348.27, variance t3 = 616.52, covariance t1–t2 = 333.87, covariance t1–t3 = 346.69 and covariance t2–t3 = 367.82.

Table 2 and Fig. 1 show the fixed effects for the OQ-45 total scale of the level-2-model:

- t1: First measurement point: The t1 estimate for the control group amounted to 84.45 OQ-45 points. This result was shown to be significantly different from zero ($T = 19.75$; $p < 0.05$). The estimate for the experimental group at t1 was 90.87 ($84.45 + 6.42$) OQ-45 points. The difference of 6.42 OQ-45 points between the experimental and the control group at t1 did not reveal any statistically significant difference ($T = 1.10$; $p = 0.28$). Thus, the psychological distress assessed with the OQ-45 in both treatment groups can be viewed as comparable at t1.
- t2: Signal measurement point: At t2, the control group achieved an estimate of 100.60 ($84.45 + 16.15$) OQ-45 points. This value was significantly higher than the OQ-45 score of the control group at t1 ($T = 10.63$; $p < 0.05$). For the experimental group, an estimate of 105.30 ($84.45 + 6.42 + 16.15 - 1.72$) OQ-45 points was found at t2. The experimental group reached 1.72 less OQ-45 points than the control group when comparing t1 to t2. This result did not attain statistical significance ($T = -0.83$; $p = 0.41$). Hence, the deterioration was significant and similar for both groups from t1 to t2.
- t3: Last measurement point: The t3 estimate for the control group amounted to 98.65 ($84.45 + 14.20$) OQ-45 points. The control group had a significantly higher OQ-45 score at t3 than at t1 ($T = 3.74$, $p < 0.05$). The estimate for the experimental group at t3 was 92.57 ($84.45 + 6.42 + 14.20 - 12.50$). When comparing differences in the estimates of both groups between t1 and t3 (estimates of the degree of deterioration), 12.50 fewer OQ-45 points were calculated for the experimental group. This difference was statistically significant ($T = -2.41$; $p < 0.05$). Therefore, it can be concluded that feedback does improve the outcome of in-patients at risk in this study.

The effect size (Hedges' g) at t1 reached $d = 0.31$. At t1, the experimental group showed more psychological distress than the control group. For calculating the effect size, the following values at t1 were used: experimental group: $M = 90.87$, $SD = 18.58$; control group: $M = 84.45$, $SD = 20.68$.

At t3, an effect size (Hedges' g) of $d = 0.23$ was found. The experimental group suffered less psychological distress than the control group at t3. The means and standard deviations at t3 used for the calculation of the effect size were: experimental group: $M = 92.57$, $SD = 25.40$; control group: $M = 98.65$, $SD = 25.46$.

Table 1

Comparison of age, gender, education, comorbidity, treatment duration and assessment points with first at-risk signal between experimental and control group.

Variable	Experimental group N = 23	Control group N = 20	Statistics
Age: M (SD)	43.45 (9.93)	47.34 (11.86)	t(41) = 1.17; p = 0.25
Gender: n (%)			p = 0.55 (FET)
	female	14 (60.9%)	10 (50.0%)
	male	9 (39.1%)	10 (50.0%)
Education: n (%)			p = 0.50 (FET)
	lower secondary school	6 (26.1%)	9 (45.0%)
	intermediate secondary school	8 (34.8%)	6 (30.0%)
	higher secondary school	6 (26.1%)	2 (10.0%)
	other type of graduation	3 (13.0%)	3 (15.0%)
Comorbidity of mental and behavioral disorders: M (SD)	2.26 (0.96)	2.65 (1.09)	t(41) = 1.24; p = 0.22
Treatment weeks from intake to discharge: M (SD)	6.22 (3.29)	5.49 (3.17)	t(41) = -0.74; p = 0.46
Treatment weeks from t1 to t3: M (SD)	5.13 (3.28)	3.95 (3.03)	t(41) = -1.22; p = 0.23
Treatment weeks from t1 to t2: M (SD)	2.65 (1.64)	1.95 (0.94)	t(41) = -1.69; p = 0.10
Treatment weeks from t2 to t3: M (SD)	2.48 (3.13)	2.00 (3.06)	t(41) = -0.51; p = 0.62
Frequencies of patients receiving the first at risk signal at different assessment points: n (%)			p = 0.55 (FET)
	Intake week + 1 week	5 (21.7%)	8 (40.0%)
	Intake week + 2 weeks	8 (34.8%)	6 (30.0%)
	Intake week + 3 weeks	6 (26.1%)	6 (30.0%)
	Intake week + 4 weeks	1 (4.3%)	-
	Intake week + 5 weeks	2 (8.7%)	-
	Intake week + 8 weeks	1 (4.3%)	-

Hence, the adjusted feedback effect occurring from t1 to t3 was $d = 0.54$ ($d = 0.31 + d = 0.23$) in favor of the experimental group.

Table 3 displays the results of the RCI for at-risk patients of the experimental and control group from t1 to t3. By providing feedback, patients at risk showed less reliable deterioration and more reliable improvement on the OQ-45 total scale. The odds ratio of reliable deterioration amounted to 0.29, indicating that reliable deterioration occurred about one-quarter (25%) to one-third (33%) less likely in the experimental group than in the control group.

Fig. 2 shows the individual change scores, OQ-45 total score t3 – OQ-45 total score t1, for all at-risk patients confirming the tendency of favoring the experimental group.

Discussion

In this study of psychosomatic in-patient treatment, 43 of 252 patients (17.1%) were at risk of treatment failure. This percentage of in-patients at risk was comparable to the percentage of at-risk patients in counseling [21]. The psychosomatically treated in-patients at risk showed a significantly improved outcome (adjusted effect size from t1 to t3: $d = 0.54$) when feedback on patient progress and CST was provided to therapists on a weekly basis. On average, the scores of patients at risk of the experimental group were only $M = 2$ OQ-45 points higher from the time of intake (t1) to the last measurement point (t3). On the other hand, patients at risk of the treatment-as-usual control group scored $M = 14$ additional OQ-45 points on average during this time period. In accordance with the results of studies pertaining to counseling [24,26,27], outpatient psychotherapy [28], and substance abuse treatment [29], this study showed that feedback on patient progress and CST can reduce the average deterioration of patients at risk in psychosomatic in-patient therapy. Therefore, feedback can have an additional effect to the common strategies for risk patients in psychosomatic in-patient treatment, e. g. intervention and supervision. Since no differences between the change patterns of the two treatment groups can be found from intake to the signal week, one can assume that the effect of weekly

feedback can only be observed after a warning signal is displayed on the feedback reports. In view of the relatively small sample size of this study, conducting more studies preferably with larger samples in order to replicate the feedback impact on patients at risk in psychosomatic in-patient treatment would seem advisable. It also should be noted, that even when feedback was provided to therapists, at t3 the in-patients at risk were worse off than at t1. Feedback on patient progress and CST only could reduce the average deterioration ($M = 2$ vs. $M = 14$). Studies investigating how in-patients at risk can be discharged in an improved condition are necessary in the future.

Despite the positive effect of feedback for psychosomatically treated in-patients at risk shown in this study, our effect size of $d = 0.24$ at t3 is lower than the results reported for a mega- and meta-analytical study based on counseling center data (computing two effect sizes: a) intent-to-treat analysis: $d = 0.70$, and b) efficacy analysis: $d = 0.44$) which compared feedback on patient progress and CST to no feedback [19]. In an other study evaluating at-risk patients in outpatient psychotherapy however, only a small effect size of $d = 0.16$ was found at t3 for feedback on patient progress and CST in comparison to treatment-as-usual [28]. Simon and colleagues suggest "that the interventions do not work as well with more disturbed patients as with the less disturbed" [28, p. 645]. Future studies analyzing the factors leading to these differences of effect sizes are required to shed light on this question. Furthermore, they could point to ways of further improving treatment. Differences between the studies pertaining to the treatment integrity could also play a role regarding the different effect sizes [38].

Considering the RCI classification of patients at risk from t1 to t3, the experimental group showed about 65% fewer reliably deteriorated patients at risk than the control group (8.7% vs. 25.0%, odds ratio: 0.29). The effect of feedback for patients at risk was further supported by the fact that the only two reliably improved at-risk patients belonged to the experimental group (13.0% vs. 0.0%). In a study on

Table 2

Fixed effects of the level-2-model (see text for details).

Parameter	Estimate OQ-45 total score	SE	df	T-statistics	p-Value
t1: Control group	84.45	4.28	43	19.75	<0.01
t1: Experimental group vs. control group	6.42	5.85	43	1.10	0.28
t1 → t2: Control group	16.15	1.52	43	10.63	<0.01
t1 → t2: Experimental group vs. control group	-1.72	2.08	43	-0.83	0.41
t1 → t3: Control group	14.20	3.80	43	3.74	<0.01
t1 → t3: Experimental group vs. control group	-12.50	5.20	43	-2.41	0.02

Abbreviations: SE = standard error, df = degrees of freedom, t1 = intake week, t2 = signal week, t3 = week of the last available OQ-45 score.

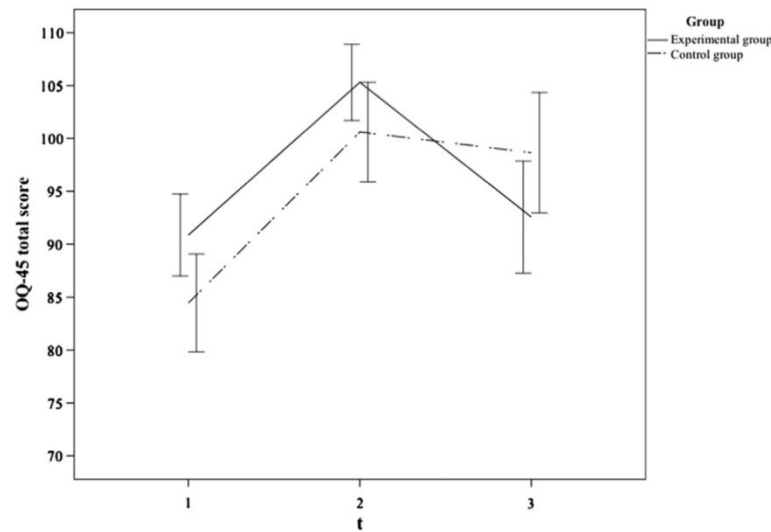


Fig. 1. OQ-45 total scale mean at t1 (intake week), t2 (signal week), and t3 (last measurement point) for experimental and control group \pm 1 SE.

outpatient psychotherapy, the number of reliably deteriorated patients could be reduced through feedback on patient progress and CST by about 50% [30]. Our results indicate that feedback could be even more powerful in reducing deterioration in in-patient than in outpatient psychotherapy. In this context, it should be noted that the current study used the more conservative RCI (21 points of change) from the German normative samples [31] and not that from the samples of the U.S. where the RCI requires only 14 or more points, thus possibly providing lower estimates of reliable change in comparison to past research.

No significant differences between the two treatment groups regarding age, gender, education, comorbidity, treatment weeks (from intake to discharge, from t1 to t3, from t1 to t2, from t2 to t3), or initial psychological dysfunction (intake OQ-45 total score) could be found. Hence, the improved outcome of the experimental group could not be attributed to significant differences in these variables between experimental and control groups. Although not statistically significant, it should be noted that the experimental group tended to be more educated, to have less comorbid disorders, and to stay longer in treatment than the control group. Furthermore, there are more patients with somatoform disorders in the control group (65.0% vs. 52.2%) who are known to improve on average less during psychotherapy than other diagnostic groups [2,39]. However, it should be taken into account that in this study the diagnoses were made by the clinic teams and not by standardized interviews, resulting in an uncertain validity of the diagnoses [40].

In psychosomatic treatment several therapists are involved in the treatment of a patient. This study showed that patients at risk have a better outcome when one therapist, the one responsible for individual

psychotherapy with the patient, is provided with weekly feedback on patient progress and CST. For further studies of psychosomatic in-patient treatment, it would also be interesting to investigate if the positive effect of feedback could be further enhanced when all therapists involved in the treatment of a patient received the feedback.

Another point for future research relates to the expected recovery curves. The expected recovery curves of the empirical algorithm used in this study are based on thousands of patients from the U.S., coming from different psychotherapy settings. Expected recovery curves based solely on psychosomatically treated in-patients could allow assessing at-risk patients more precisely by comparing individual patient progress against the expected progress in psychosomatic in-patient treatment. Nevertheless, in this study the expected recovery curves from the U.S. showed predictive power in German psychosomatic in-patient treatment: When in-patients were first classified as at-risk in the signal week (t2), both groups – on average – deteriorated significantly compared to intake (t1). Furthermore, 7 of altogether 8 patients who deteriorated at t3 were labeled as patients at risk. This hit rate of 87.5% was comparable to the results from the U.S. [12,13]. In addition, the 7 at-risk patients with deterioration (87.5% of all deteriorated patients) were detected during the first four treatment weeks (intake week + 3 weeks). In the U.S., 85% of all patients who deteriorated were identified during the first four weeks (intake interview + 3 sessions) [12]. Early feedback on risk of deterioration is essential, since the later at-risk patients are detected the less opportunities therapists have to take action based on the feedback. On average, the first warning signal appeared in the feedback reports $M=2.65$ (experimental group) and $M=1.95$ (control groups) weeks after the intake week.

Another factor to consider in this study is the applicability of the ASC (CST instrument) in psychosomatic in-patient setting, where the social support scale may not be that useful as in outpatient psychotherapy, since in-patients usually do not have that much contact to the family or friends. Accordingly, further analysis of our ASC data should concentrate on evaluating the applicability in psychosomatic in-patient settings.

To summarize – if the results of this study are replicated – monitoring and feedback in psychosomatic in-patient treatment should be applied in routine care, since it can improve the outcome of patients at risk of

Table 3
Reliable change index (RCI) classification (t1 to t3).

RCI	Experimental group N = 23	Control group N = 20
Reliably improved (improved by at least 21 OQ-45 points)	3 (13.0%)	0 (0.0%)
Unchanged (fewer than 21 OQ-45 points change)	18 (78.3%)	15 (75.0%)
Reliably deteriorated (worse by at least 21 OQ-45 points)	2 (8.7%)	5 (25.0%)

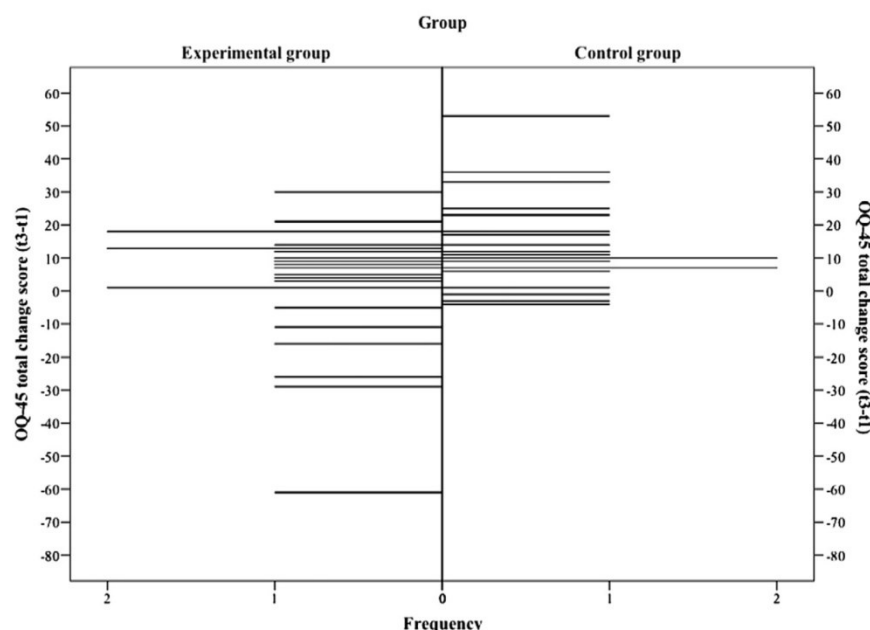


Fig. 2. Frequencies of individual OQ-45 total change scores (t3–t1) for all patients at-risk. Positive values indicate deterioration. Negative values indicate improvement.

treatment failure. The Ethics Code of the APA advising that “Psychologists take reasonable steps to avoid harming their clients/patients... and to minimize harm where it is foreseeable and unavoidable” [41, p.1065] strongly supports this position.

Conflict of interest

The authors report no financial or other relationships relevant to the subject of this article.

Acknowledgments

This research was funded by the Susa Young Gates University Professorship awarded to Michael J. Lambert.

References

- [1] Zaudig M. Fifty years of psychosomatic inpatient treatment. *Z Psychosom Med Psychother* 2004;50:355–75.
- [2] Steffanowski A, Löschmann C, Schmidt J, Wittmann WW, Nübling R. Metaanalyse der Effekte psychosomatischer Rehabilitation. Bern: Huber; 2007.
- [3] Tritt K, von Heymann F, Loew TH, Benker B, Bleichner F, Buchmüller R, et al. Patients of a psychotherapeutic inpatient setting: patient description and effectiveness of treatment. *Psychother Psychol Psychother Med Klin Psychol* 2003;8: 244–51.
- [4] Probst T, von Heymann F, Zaudig M, Konermann J, Lahmann C, Loew T, et al. The effectiveness of psychosomatic inpatient treatment – results of a multicenter follow-up study. *Z Psychosom Med Psychother* 2009;55:409–20.
- [5] Dilcher K, Mestel R, Klingelhöfer J, Köbel W, Sprenger B, Stauss K. Psychosomatische Kliniken. In: Gerdes N, Weidemann H, Jäckel WH, editors. *Die Protos-Studie*. Darmstadt: Steinkopff; 2000. p. 173–202.
- [6] Franz M, Janssen P, Lentsche H, Schmidtke V, Tetzlaff M, Martin K, et al. On the effect of psychoanalytic oriented psychotherapy – an inpatient multicenter study. *Z Psychosom Med Psychother* 2000;46:242–58.
- [7] Bergin AE. The effects of psychotherapy: negative results revisited. *J Couns Psychol* 1963;10:244–50.
- [8] Mohr DC. Negative outcome in psychotherapy: a critical review. *Clin Psychol Sci Pract* 1995;2:1–27.
- [9] Lilienfeld SO. Psychological treatments that cause harm. *Perspect Psychol Sci* 2007;2:53–70.
- [10] Tritt K, Bidmon RK, von Heymann F, Joraschky P, Lahmann C, Nickel MK, et al. An approach to psychotherapeutic health service research: ten theses. *Psychother Psychol Psychother Med Klin Psychol* 2007;12:47–59.
- [11] Finch AE, Lambert MJ, Schaale BG. Psychotherapy quality control: the statistical generation of expected recovery curves for integration into an early warning system. *Clin Psychol Psychother* 2001;8:231–42.
- [12] Lambert MJ, Whipple JL, Bishop MJ, Vermeersch DA, Gray GV, Finch AE. Comparison of empirically derived and rationally derived methods for identifying clients at risk for treatment failure. *Clin Psychol Psychother* 2002;9:149–64.
- [13] Spielmanns GI, Masters KS, Lambert MJ. A comparison of rational versus empirical methods in prediction of negative psychotherapy outcome. *Clin Psychol Psychother* 2006;13:202–14.
- [14] Lambert MJ, Kahler M, Harmon C, Shimokawa K, Burlingame G. Administration and Scoring Manual for the Outcome Questionnaire-45.2. Orem, UT: American Professional Credentialing Services; 2011.
- [15] Hannan C, Lambert MJ, Harmon C, Nielsen SL, Smart DW, Shimokawa K, et al. A lab test and algorithms for identifying clients at risk for treatment failure. *J Clin Psychol* 2005;61:155–63.
- [16] Lambert MJ. Helping clinicians to use and learn from research-based systems: the OQ-analyst. *Psychotherapy* 2012;49:109–14.
- [17] National Registry of Evidence-based Programs and Practices (NREPP). SAMHSA's National Review of Evidence-Based Practices – OQ-analyst. OQ@-Analyst. Salt Lake City, UT: OQ Measures; 2009 [www.oqmeasures.com].
- [18] Howard KI, Moras K, Brill PL, Martinovich Z, Lutz W. Evaluation of psychotherapy: efficacy, effectiveness, and patient progress. *Am Psychol* 1996;51:1059–64.
- [19] Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *J Consult Clin Psychol* 2010;78:298–311.
- [20] Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins EJ. The effects of providing therapists with feedback on patient progress during psychotherapy: are outcomes enhanced? *Psychother Res* 2001;11:49–68.
- [21] Lambert MJ, Whipple JL, Vermeersch DA, Smart DW, Hawkins EJ, Nielsen SL, et al. Enhancing psychotherapy outcomes via providing feedback on client progress: a replication. *Clin Psychol Psychother* 2002;9:91–103.
- [22] Hawkins EJ, Lambert MJ, Vermeersch DA, Slade KL, Tuttle KC. The therapeutic effects of providing patient progress information to therapists and patients. *Psychother Res* 2004;14:308–27.
- [23] Newnham EA, Hooke GR, Page AC. Progress monitoring and feedback in psychiatric care reduces depressive symptoms. *J Affect Disord* 2010;127:139–46.
- [24] Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ. Improving the effects of psychotherapy: the use of early identification of treatment and problem-solving strategies in routine practice. *J Couns Psychol* 2003;50: 59–68.
- [25] Lambert MJ, Bailey R, Kimball K, Shimokawa K, Harmon SC, Slade K. Clinical support tools manual-brief version-40. Salt Lake City: OQ Measures; 2007.

- [26] Harmon S, Lambert MJ, Smart DM, Hawkins E, Nielsen SL, Slade K, et al. Enhancing outcome for potential treatment failures: therapist–client feedback and clinical support tools. *Psychother Res* 2007;17:379–92.
- [27] Slade K, Lambert MJ, Harmon S, Smart DW, Bailey R. Improving psychotherapy outcome: the use of immediate electronic feedback and revised clinical support tools. *Clin Psychol Psychother* 2008;15:287–303.
- [28] Simon W, Lambert MJ, Harris MW, Busath G, Vazquez A. Providing patient progress information and clinical support tools to therapists: effects on patients at risk of treatment failure. *Psychother Res* 2012;22:638–47.
- [29] Crits-Christoph P, Ring-Kurtz S, Hamilton JL, Lambert MJ, Gallop R, McClure B, et al. A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *J Subst Abuse Treat* 2012;42:301–9.
- [30] Simon W, Lambert MJ, Busath G, Vazquez A, Berkeljon A, Hyer K, et al. Effects of providing patient progress feedback and clinical support tools to psychotherapists in an inpatient eating disorders treatment program: a randomized controlled study. *Psychother Res* 2013;23:287–300.
- [31] Lambert MJ, Hannover W, Nisslmüller K, Richard M, Kordy H. Questionnaire on the results of psychotherapy: reliability and validity of the German translation of the Outcome Questionnaire 45.2 (OQ-45.2). *Z Klin Psychol Psychother* 2002;31:40–7.
- [32] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for Patient-Reported Outcomes (PRO) measures. Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005;8:94–104.
- [33] Singer JD, Willett JB. *Applied longitudinal data analysis. Modeling Change and Event Occurrence*. London: Oxford University Press; 2003.
- [34] Gollner R, Gollwitzer M, Heider J, Zaby A, Schröder A. Analysis of longitudinal data with hierarchical linear models. *Z Klin Psychol Psychother* 2010;39:179–88.
- [35] Durlak JA. How to select, calculate, and interpret effect sizes. *J Pediatr Psychol* 2009;34:917–28.
- [36] Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991;59:12–9.
- [37] World Health Organization WHO. *ICD-10: The ICD-10 Classification of Mental and Behavioural Disorders*. Geneva: WHO; 1992.
- [38] Sechrest L, Phillips MA, Redner R, Yeaton WH. Some neglected problems in evaluation research: strength and integrity of treatments. In: Sechrest L, West SG, Phillips M, Redner R, Yeaton WH, editors. *Evaluation studies review annual*, vol. 4. Beverly Hills, CA: Sage; 1979. p. 15–35.
- [39] Huber D, Albrecht C, Hautum A, Henrich G, Klug G. Effectiveness of inpatient psychodynamic psychotherapy: a follow-up study. *Z Psychosom Med Psychother* 2009;55:189–99.
- [40] Knappe S, Runge J, Beesdo K, Jacobi F, Wittchen H-U. Diagnosing mental disorders: gold or tin standard? Critical comments on standardized diagnostic interviews and clinical routine diagnoses. *PPmP* 2008;58:72–5.
- [41] American Psychological Association (APA). *Ethical principles of psychologists and code of conduct*. *Am Psychol* 2002;57:1060–73.

Corrigendum to correct errors in the Probst et al. (2013) publication (sent to the Editor of the Journal on 03/24/2014):

1) *Text with citation errors in the introduction on pages 255-256:* While patients not labeled as at-risk patients have deterioration rates between 0.3% and 1.3%, at-risk patients show deterioration rates between 8.4% and 19.4% (results for the empirical algorithm [12,13]).

Corrected text: While patients not labeled as at-risk patients have deterioration rates between 0.0% and 1.3%, at-risk patients show deterioration rates between 19.4% and 30.5% (results for the empirical algorithm [12,13]).

2) *Text with citation errors in the introduction on page 256:* Providing patient progress feedback to therapists has been shown to improve the outcome of patients at risk in counseling [19,20], outpatient psychotherapy [21], and psychiatric care [22]. At-risk patients whose therapists received patient progress feedback also stayed longer in treatment [23], probably contributing to the improved outcome.

Corrected text: Providing patient progress feedback to therapists has been shown to improve the outcome of patients at risk [19] in counseling [20,21] outpatient psychotherapy [on a trend level 22], and psychiatric care [23]. At-risk patients whose therapists received patient progress feedback also stayed longer in treatment [20,21], probably contributing to the improved outcome.

3) *Text with interpretation errors in the results on page 258:* The odds ratio of reliable deterioration amounted to 0.29, indicating that reliable deterioration occurred about one-quarter (25%) to one-third (33%) less likely in the experimental group than in the control group.

Corrected text The odds ratio of reliable deterioration amounted to 0.29, indicating that the odds for reliable deterioration at t3 were approximately 3.5 times higher in the control group than in the experimental group.

4) *Text with citation errors in the discussion on page 258:* ... the results reported for a mega- and metaanalytical study based on counseling center data (computing two effect sizes: a) intent-to-treat analysis: $d = 0.70$, and b) efficacy analysis: $d = 0.44$) which compared feedback on patient progress and CST to no feedback [19].

Corrected text: ... the results reported for a mega- and metaanalytical study based on counseling center data (computing two effect sizes: a) intent-to-treat analysis: $d = 0.44$, and b) efficacy analysis: $d = 0.70$) which compared feedback on patient progress and CST to no feedback [19].

III. 2. Publikation 2

Probst, T., Lambert, M. J., Dahlbender, R. W., Loew, T. H., & Tritt, K. (2014). Providing patient progress feedback and clinical support tools to therapists: Is the therapeutic process of patients on-track to recovery enhanced in psychosomatic in-patient therapy under the conditions of routine practice? *Journal of Psychosomatic Research*, 76, 477-484.

doi: 10.1016/j.jpsychores.2014.03.010.

Link: <http://www.sciencedirect.com/science/article/pii/S0022399914000798>



Providing patient progress feedback and clinical support tools to therapists: Is the therapeutic process of patients on-track to recovery enhanced in psychosomatic in-patient therapy under the conditions of routine practice?[☆]



Thomas Probst^{a,*}, Michael J. Lambert^b, Reiner W. Dahlbender^c, Thomas H. Loew^d, Karin Tritt^d

^a University of Regensburg, Department of Psychology, Clinical Psychology and Psychotherapy, Germany

^b Brigham Young University, Department of Psychology, United States

^c University Clinic of Ulm, Clinic for Psychosomatic Medicine and Psychotherapy, Germany

^d University Clinic of Regensburg, Department of Psychosomatics, Germany

ARTICLE INFO

Article history:

Received 19 October 2013

Received in revised form 19 March 2014

Accepted 19 March 2014

Keywords:

Inpatient treatment
Therapeutic process
Psychosomatics
Psychotherapy
Routine clinical care
Feedback
Clinical support Tools

ABSTRACT

Objectives: In previous studies of patients on-track to recovery (OT) involving therapists receiving only patient progress feedback without clinical support tools (CST) inconsistent results were found. Possible effects of combining patient progress feedback with CST on OT patients remain unclear.

Methods: At intake (t1), 252 patients of two in-patient psychosomatic clinics were randomized either into the experimental group (EG) or the treatment-as-usual control group (CG). Both groups were monitored weekly using the self-report instruments "Outcome Questionnaire" (OQ-45) and "Assessment of Signal Cases" (ASC). Therapists received weekly patient progress feedback (OQ-45) and CST feedback (ASC) only for EG patients starting at the week following intake (t2). Patients who did not deviate negatively from expected recovery curves by at least one standard deviation were considered OT patients ($N = 209$; $N_{EG} = 111$; $N_{CG} = 98$). Since therapists received feedback at t2 for the first time, different patterns of change (OQ-45 scales) between the groups from t1 to t2, t2 to t3 (intake + two weeks), t2 to t4 (intake + three weeks), and t2 to t5 (last available OQ-45 score) were evaluated by multilevel models.

Results: Merely from t2 to t3, the EG improved significantly more on the OQ-45 symptom distress scale than the CG ($p < 0.05$; $g = 0.12$).

Conclusion: Providing patient progress feedback and CST to therapists did not substantially surpass treatment-as-usual for OT patients in this explorative study except for a very small time-limited enhancement of symptom change.

© 2014 Elsevier Inc. All rights reserved.

Introduction

An important future task of psychotherapy research is to develop and to implement methods that can help therapists improve ongoing psychotherapies [1]. The study of ongoing psychotherapies is linked to patient focused research [2]. In contrast to treatment focused research, which analyzes the outcome of different treatment groups to identify empirically supported treatments, patient-focused research focuses on the individual patient's progress throughout the course of therapy. Frequent assessments (e.g. session by session, weekly, or even daily) of the

patient's problems enable tracking the patient's response to treatment. In tracking, one compares the individual progress of a given patient against the progress that would be expected for this patient in order to classify the patient's response. Two different types of algorithms representing such "expected recovery curves" are available for the self-report "Outcome Questionnaire" (OQ-45) [3]: Rational and empirical algorithms. The expected recovery curves of the rational algorithm are based on expert ratings pertaining to what progress should be achieved during which time period of psychotherapy [4,5]. The expected recovery curves of the empirical algorithm are based on the data of 11,492 patients treated in different psychotherapeutic settings in the U.S. [6]: For each individual patient, the empirical algorithm provides an expected recovery curve based upon at least 220 patients with an initial intake OQ-45 score comparable to the patient in question [6]. The expected recovery curves of the empirical algorithm depend only on the initial OQ-45 score, because other variables (e.g. diagnosis, client demographics or therapists' orientation) were shown to be responsible

[☆] The work was conducted at the Psychosomatics Department of the Hospital in Donaustauf, Germany, and the Psychosomatic Hospital "Am Schönen Moos" in Bad Saulgau, Germany.

* Corresponding author at: University of Regensburg, Clinical Psychology and Psychotherapy, 93051 Regensburg, Germany. Tel.: +49 941 6309158 54.

E-mail address: thomas.probst@psychologie.uni-regensburg.de (T. Probst).

for less than 1% of the outcome variance after taking the initial score plus the change occurring between session 1 and session 3 into account [1].

By tracking and comparing the progress of individual patients against expected recovery curves, patients at risk of treatment failure can be identified quite early in the course of therapy. These patients are also called not-on-track patients (NOT) since they deviate notably (as defined by the algorithms) from expected recovery curves. These NOT patients have a significantly higher risk of leaving psychotherapy deteriorated than patients on-track to recovery (OT), whose progress lies within the boundaries of the expected recovery curves [4,5]. The software OQ-Analyst [7] – rated as an evidence-based practice with excellent reviews for its training material and easy implementation [8] – was designed for monitoring the patient's progress by weekly OQ-45 assessments and for comparing the individual patient's progress against the expected recovery curve. Furthermore, the OQ-Analyst automatically transforms the OQ-45 data into feedback reports for therapists, which include a graph showing the individual's progress compared to the expected progress (*patient progress feedback*). Clinical support tools (CST) [9] are also integrated into the OQ-Analyst. CST are intended to help therapists identify and then focus on domains that are thought to have influenced the given patient's progress (e.g., therapeutic alliance). Furthermore, when problems in these domains occur, CST are intended to assist therapists in solving these problems.

The CST of the OQ-Analyst consist of two components: The first CST component is the self-report questionnaire "Assessment of Signal Cases" (ASC) [9] measuring the following domains that are considered to influence patient progress: Therapeutic alliance, motivation, social support, and life events. The ASC was initially designed for NOT patients, but it could also be used for all patients in order to track these domains on a weekly basis throughout therapy. Using the OQ-Analyst, feedback reports of the patient's results on the ASC scales are available for therapists, giving them indications of existing problems of the specific patient regarding the ASC domains (*CST feedback*). The second CST component is a decision tree for handling these problems assessed by the ASC.

Studies on providing patient progress feedback and CST to therapists in counseling have shown that 1) patient progress feedback is superior to treatment-as-usual for NOT patients and 2) adding CST to patient progress feedback surpasses patient progress feedback alone as well as treatment-as-usual for NOT patients [10–12]. Further research in outpatient psychotherapy [13], substance abuse treatment [14], and psychosomatic in-patient therapy [15] found that treatment-as-usual can be enhanced for NOT patients when therapists receive patient progress feedback and CST. Regarding OT patients, patient progress feedback alone (without CST) was compared against treatment-as-usual in five previous studies [10,11,16–18]. In two randomized controlled trials (RCT) and in one quasi-experimental study, the outcome of OT patients was not significantly affected by patient progress feedback [10,16,17]. However, another quasi-experimental study with archival control data found that providing feedback on patient progress to therapists significantly improves the outcome of OT patients [11]. In addition, a mega-analysis based on the combined data of all studies using the above mentioned algorithms of the OQ-45 to classify patients as OT/NOT suggested that patient progress feedback surpasses treatment-as-usual in producing positive outcomes for OT cases [18]: In this mega-analysis, two different effect sizes (g) were reported: a) intent-to-treat analysis: $g = 0.12$ (2.8 points difference on the OQ-45 total scale); and b) efficacy analysis: $g = 0.30$ (6.5 points difference on the OQ-45 total scale). As the previous studies only applied patient progress feedback without CST for OT patients, possible effects of combining patient progress feedback with CST on OT patients remain unclear. In this project, patient progress feedback was combined for the first time with CST for OT patients to investigate whether this combined feedback is superior to treatment-as-usual not only for NOT patients, but also for OT patients. In an explorative study, we examined whether this combined feedback can enhance the therapeutic process of

psychosomatically treated OT in-patients regarding global and/or specific psychological problems.

Methods

Questionnaires

Outcome questionnaire (OQ-45)

The German version of the self-report questionnaire OQ-45 [3,19] was used to track patient progress. All 45 items are assessed on a five-point Likert scale and refer to the past week. The OQ-45 provides one total score (45 items; range: 0–180) and the following three subscales: Symptom distress scale (25 items; range: 0–100), interpersonal relations scale (11 items; range: 0–44), and social role performance scale (9 items; range: 0–36). The higher the values on the OQ-45 scales, the worse the impairment. Lambert et al. [19] reported the following internal consistency coefficients (α) and test–retest reliability coefficients (r) for the German version of the OQ-45: Total scale ($\alpha = 0.93$; $r = 0.88$), symptom distress scale ($\alpha = 0.90$; $r = 0.87$), interpersonal relations scale ($\alpha = 0.81$; $r = 0.81$), and social role performance scale ($\alpha = 0.59$; $r = 0.71$). The following intercorrelations (r) between the scales of the OQ-45 were reported [19]: $r = 0.95$ (symptom distress scale vs. total scale), $r = 0.66$ (symptom distress scale vs. social role performance scale), $r = 0.66$ (symptom distress scale vs. interpersonal relations scale), $r = 0.83$ (interpersonal relations scale vs. total scale), $r = 0.59$ (interpersonal relations scale vs. social role performance scale), and $r = 0.79$ (social role performance scale vs. total scale).

Assessment of signal cases (ASC)

The self-report questionnaire ASC [9] was the CST instrument used in this study. It consists of 40 items rated on a five-point Likert scale which refer to the previous week. The ASC includes four scales, which are viewed as influencing therapeutic progress [9]: Therapeutic alliance (11 items), social support (11 items), motivation (9 items), and life events (9 items). The ASC was translated into German via back-translation method for this study. Before the study commenced, cognitive debriefing interviews were done with ten in-patients to test the quality of the translation [20]. In this study, the CST instrument ASC was used for all patients as an instrument to monitor the course of therapeutic alliance, social support, motivation, and life events throughout therapy – and not only administered to NOT patients like in the previous studies [10–14]. The reliability of the four scales of the German ASC version was tested with the internal consistency coefficient Cronbach's α . The following coefficients for intake ASCs were found for all patients (not only OT patients): Therapeutic alliance scale: $\alpha = 0.89$; social support scale: $\alpha = 0.76$; motivation scale: $\alpha = 0.78$; and life events scale: $\alpha = 0.71$.

Study design

The Psychosomatics Department of the Hospital in Donaustauf, Germany and the Psychosomatic Hospital "Am Schönen Moos" in Bad Saulgau, Germany participated in the study. Both clinics provided an integrative multi-modal and multi-method psychosomatic in-patient treatment of more than 24 treatment hours per week, including verbal therapy (individual and group psychotherapy), extra-verbal therapy (art, dance, and music), relaxation, and mindfulness training. In both clinics multi-professional teams (psychologists, physicians, nurses, etc.) worked with the patients. All patients in Donaustauf were encouraged to take part in the study which lasted from 10/01/2010 through 07/06/2012; while in Bad Saulgau only patients with private health care were asked to participate in the study which took place from 01/16/2012 through 04/06/2012. At intake, all patients were randomized either into the experimental group (EG) or the treatment-as-usual control group (CG). Over the course of their stay the patients were asked to fill out the German paper and pencil versions of the OQ-45 and the ASC

once every week anytime between Friday and Sunday. While both groups (EG and CG) were monitored on a weekly basis with the OQ-45 and the ASC, feedback on both instruments was provided only for EG patients. Furthermore, only the individual therapists of those EG patients received feedback.

The feedback reports generated by the OQ-Analyst (patient progress feedback as well as CST feedback) were given to the therapists of all EG patients in a closed envelope on Monday, after the OQ-45 and ASC data had been entered into the OQ-Analyst. The reports of a given patient displayed the patient's OQ-45 progress against the expected recovery curve as defined by the empirical algorithm of the OQ-45 (patient progress feedback) as well as the results of the ASC scales (CST feedback).

Since at least two OQ-45 assessments are necessary to measure patient progress, the patient progress feedback plus CST feedback was given to therapists on a weekly basis from the second assessment point on. Before the study started, the therapists were briefed in detail on how to interpret the feedback reports and the therapists received the decision tree of the OQ-Analyst for handling problems in the domains assessed with the ASC. No experimental control was exercised on the therapists regarding their use of the feedback reports, i.e. they were allowed to discuss the results with the patients, the clinic team, and/or their supervisors. We decided to grant the participating therapists this freedom, assuming that any sort of punitive measures and control mechanisms employed could influence the participants negatively. In this paper, only OT patients – according to the empirical algorithm of the OQ-45 [6] – were analyzed. A given patient was classified as an OT patient, when the patient did not – at any OQ-45 assessment during psychosomatic treatment – deviate negatively from his or her expected recovery curve by one standard deviation or more. Those patients who deviated negatively by at least one standard deviation from the expected recovery curves were considered as NOT patients; the feedback effect on psychosomatically treated NOT in-patients was reported elsewhere [15]. For NOT patients, therapists received a warning signal (yellow or red signal) in the feedback reports. For OT patients, a green signal or a white signal (patient's score is located below the OQ-45 cut-off dividing the functional and dysfunctional group) was provided to therapists via the feedback reports.

Since the intake OQ-45 score functions as the baseline for the expected recovery curves of the empirical OQ-45 algorithm and generating feedback on patient progress requires at least two assessment points, effects of feedback can only be studied from the third assessment point on. Therefore, only those OT patients were included in the statistical analyses of feedback effects who had an available intake OQ-45 score and who had filled out the OQ-45 at least at two more assessment points.

Statistical analysis

SPSS 21.0 was used for the statistical analyses applying a significance level of $p \leq 0.05$. All tests were performed two-tailed. Due to the explorative nature of this study, p -values uncorrected for multiple testing will be reported in this paper. Means (M) and standard deviations (SD) were calculated for the sample description. Comparisons between patients meeting the inclusion criteria (as defined above) and excluded patients were done using Fishers' exact tests (FET) for nominal and independent t -tests for metric variables. The same statistical tests were used to test for differences between the two treatment groups at intake. To analyze the impact of feedback on the therapeutic process of psychosomatically treated OT patients, multilevel models for discontinuous change [21: chapter 6, 22: model 4] using maximum likelihood estimation (ML) were applied. For each OQ-45 scale, a multilevel model for discontinuous change was used with the mixed command in SPSS 21.0. Within the repeated subcommand, an unstructured variance-covariance structure was selected for each model. The multilevel models used in this study were level-2-models: Assessment points as level 1, patients as level 2, and treatment group as a predictor on level 2. For the

discontinuous level-2-models of change, the following five assessment points were chosen:

- t1 Intake week.
- t2 Intake week + 1 week: At t2, the first feedback reports (patient progress and CST) were provided to therapists for their EG patients.
- t3 Intake week + 2 weeks.
- t4 Intake week + 3 weeks.
- t5 Discharge week or last available OQ-45 assessment. As in previous studies [11,12,15], the last observation carried forward method (LOCF) was applied to obtain an estimation of the feedback effect at time of post-intervention which can be compared with the results of those previous studies. On average, t5 assessment was done at $t1 + M = 3.62$ ($SD = 2.63$) weeks for the EG and at $t1 + M = 3.43$ ($SD = 1.53$) weeks for the CG.

As described by Göllner et al. [22], the five assessment points were coded into four contrast variables. The assessment point t2 was taken as the reference, since therapists received their first feedback reports for their EG patients at t2 (start of the intervention). Simple contrast tests were used to evaluate different patterns of change in the OQ-45 scales between the CG and the EG from t1 to t2, t2 to t3, t3 to t4, as well as from t2 to t5.

We did not add further measurement points (e.g. intake week + 4 weeks) to the multilevel models, since OQ-45 data was not available for the majority of the included 252 patients at these assessment points due to the short treatment durations in the clinics: For the measurement point "intake + 4 weeks" missing data rises to 73.47% for the CG and to 72.97% for the EG. Information on the percentage of missing data for the included assessment points is shown in Fig. 1.

In this analysis, multilevel models were chosen instead of traditional repeated analyses of variance (rANOVA), because the assumption of the rANOVA regarding equidistant measurement points was only met for measurement points t1, t2, t3, and t4 ($t_x + 1 = t_x + 1$ week), but not for t5 (mean assessment of $t5 = t4 + 0.62$ weeks for the EG and $t5 = t4 + 0.43$ weeks for the CG). Moreover, multilevel models allow handling missing data more flexibly: In a rANOVA with the five above mentioned assessment points, only patients with no missing data would be included, whereas in a multilevel model all patients meeting the above mentioned inclusion criteria to study the feedback effects can be analyzed.

We chose models for discontinuous change instead of models for linear change, since weekly changes in psychosomatic in-patient therapy do not seem to proceed in a linear fashion: As has been shown [23], the improvement rates occurring during the week after the intake week tend to be substantially higher than the improvement rates reached during the following weeks.

We also computed an adjusted effect size Hedges g [24] when the multilevel models revealed a statistically significant change between the groups between t2 and any other measurement point: The effect size at t2 was added to the effect size of the measurement point, at which the EG differed significantly from the CG. For calculation of the effect sizes, the estimations of the multilevel models were used.

Sample

Of all 436 patients registered during the period of the study, 252 patients from the two clinics met the inclusion criteria: Available OQ-45 scores for the intake week and for at least two more weeks. Fig. 1 depicts the flow chart of the total sample of our feedback project, including the attrition rate of the sample of this study.

The results of the comparisons between included and excluded patients regarding socio-demographic and diagnostic characteristics (drop-out analysis) are presented in Table 1. The included patients did not differ significantly from the excluded patients regarding the variables gender, age, and comorbidity (amount of mental and behavioral diagnoses). Regarding education, a significant difference was found:

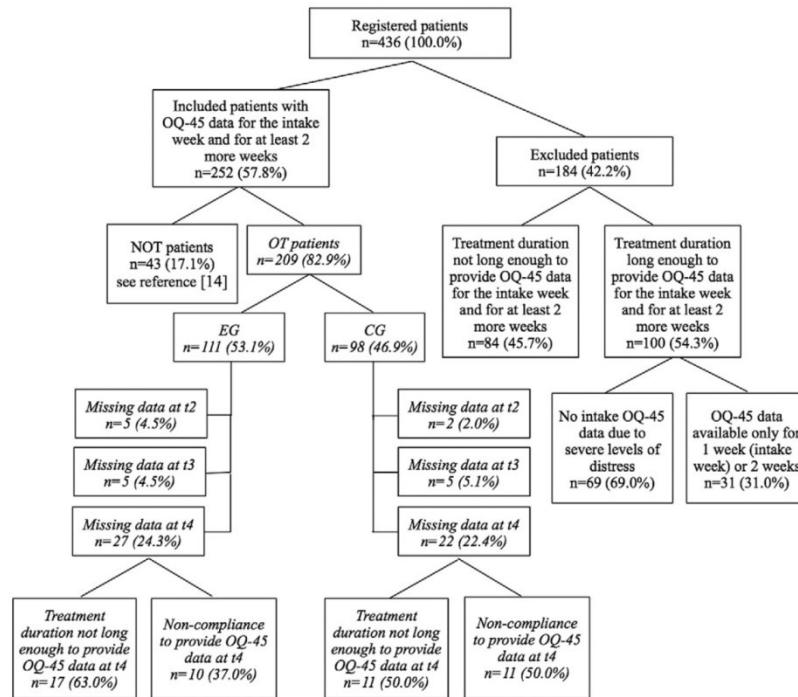


Fig. 1. Flow chart of the total sample of this feedback project in psychosomatic in-patient therapy. Data of the patients analyzed in this paper is reported in *italics*. Abbreviations: OQ-45 = "Outcome Questionnaire", NOT = not-on-track, OT = on-track, EG = experimental group, CG = control group, t2 = intake week + 1 week, t3 = intake week + 2 weeks, t4 = intake week + 3 weeks.

The excluded patients gave more often no information on their level of education than the included patients.

Applying the empirical algorithm of the OQ-45, 209 of the included 252 patients were classified as OT (82.94%) and 43 patients were considered as NOT (17.06%). Of all 209 OT patients, 111 patients were participants of the EG (53.11%) and 98 patients belonged to the CG (46.89%). In one clinic, 50.00% of the OT patients belonged to the EG/CG. In the other clinic, 53.33% of the OT patients were participants of the EG and 46.67% belonged to the CG.

Altogether, 17 therapists – who worked with the 209 OT patients in individual therapy – participated in the study. Due to the a priori confidentiality agreement with the clinics and therapists, no differences between the clinics and therapists will be reported.

Considering all comorbid diagnoses of the OT patients, the most frequent diagnoses according to Chapter V (mental and behavioral disorders) of the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10) [25] included

depressive disorders F32 and F33 (64.59% of all OT patients; 62.16% of the EG; 67.35% of the CG), somatoform disorders F45 (58.85% of all OT patients; 58.56% of the EG; 59.18% of the CG), and anxiety disorders F40 and F41 (26.32% of all OT patients; 23.42% of the EG; 29.59% of the CG). The diagnoses were made by the clinic teams.

Table 2 displays age, gender, education, comorbidity (amount of mental and behavioral diagnoses), intake scores of the OQ-45 scales, and treatment duration of the OT patients. No significant differences between the EG and the CG were found on these variables.

Results

The fixed effects of the level-2-models are presented in Tables 3–6. The results of the unstructured variance-covariance matrices were uploaded as supplementary material to the paper, accessible through the website of the journal.

Table 3 displays the results for the OQ-45 total scale. The improvement between t1 and t2, t2 and all other measurement times was statistically significant for the CG. The

Table 1
Comparisons between included and excluded patients regarding socio-demographic and diagnostic variables.

Variable	Included patients N = 252	Excluded patients N = 184	Statistics
Age: M (SD)	47.62 (13.44)	49.63 (14.60)	t(434) = -1.49; p = 0.14
Gender: n (%)			p = 0.49 (FET)
	Female	149 (59.13%)	115 (62.50%)
	Male	103 (40.87%)	69 (37.50%)
Education: n (%)			p < 0.01 (FET)
	Lower secondary school	103 (40.87%)	79 (42.93%)
	Intermediate secondary school	73 (28.97%)	44 (23.91%)
	Higher secondary school	54 (21.43%)	34 (18.48%)
	Other type of graduation	22 (8.73%)	13 (7.07%)
	No available information on education	0 (0.00%)	14 (7.61%)
Comorbidity (number of mental and behavioral diagnosis): M (SD)	2.28 (1.05)	2.19 (1.08)	t(434) = 0.89; p = 0.38

Table 2

Comparisons between the experimental group (EG) and the control group (CG) regarding gender, age, education, comorbidity, scales of the "Outcome Questionnaire" (OQ-45) at t1 (intake week), and treatment duration.

Variable	EG N = 111	CG N = 98	Statistics
Age: M (SD)	49.09 (12.83)	47.33 (14.24)	t(207) = -0.94; p = 0.35
Gender: n (%)			p = 0.89 (FET)
	Female	58 (59.18%)	
	Male	40 (40.82%)	
Education: n (%)			p = 0.07 (FET)
	Lower secondary school	40 (40.82%)	
	Intermediate secondary school	33 (33.67%)	
	Higher secondary school	22 (22.45%)	
	Other type of graduation	3 (3.06%)	
Comorbidity (number of mental and behavioral diagnoses): M (SD)	2.20 (1.12)	2.31 (0.99)	t(207) = 0.74; p = 0.46
OQ-45 total scale at t1: M (SD)	81.71 (26.03)	80.22 (25.56)	t(207) = -0.42; p = 0.68
OQ-45 symptom distress scale at t1: M (SD)	50.68 (16.47)	49.71 (16.11)	t(207) = -0.43; p = 0.67
OQ-45 interpersonal relations scale at t1: M (SD)	15.32 (7.36)	15.47 (7.03)	t(207) = 0.15; p = 0.89
OQ-45 social performance scale at t1: M (SD)	15.71 (5.85)	15.04 (5.97)	t(207) = -0.82; p = 0.41
Treatment weeks from intake to discharge: M (SD)	4.87 (3.03)	4.68 (1.68)	t(207) = -0.53; p = 0.60
Treatment weeks from t1 to t5: M (SD)	3.62 (2.63)	3.43 (1.53)	t(207) = -0.64; p = 0.53

EG did not show a significantly different course in comparison to the CG in any of the investigated weeks on the OQ-45 total scale.

The results for the OQ-45 symptom distress scale are shown in Table 4. As there was a significant feedback effect on the OQ-45 symptom distress scale between t2 and t3, the symptom improvement for both groups is displayed in Fig. 2 and described in more detail:

As Table 4 shows, at t2, the CG scored 44.65 points on the OQ-45 symptom distress scale. This result was shown to be significantly different from zero ($T = 26.01$; $p < 0.01$). Compared to the CG, the EG scored 1.37 points higher at t2 ($44.65 + 1.37 = 46.02$). This group difference did not attain statistical significance ($T = 0.58$; $p = 0.56$). Whereas the symptoms of the CG improved significantly by 5.07 points between t1 and t2 ($T = 6.43$; $p < 0.01$), the EG improved 0.41 points less on this scale ($5.07 - 0.41 = 4.66$). Nevertheless, this observed difference did not achieve statistical significance ($T = -0.38$; $p = 0.71$). From t2 to t3, the CG again improved significantly 1.56 points on the OQ-45 symptom distress

scale ($T = -2.23$; $p = 0.03$). After therapists received the first patient progress feedback and CST feedback for their EG patients at t2, the symptoms of the EG improved significantly ($T = -2.20$; $p = 0.03$) 2.12 points more than the symptoms of the CG during this treatment phase ($1.56 + 2.12 = 3.68$). The corresponding adjusted effect size amounted to $g = 0.12$ ($g_{t2} = 0.08 + g_{t3} = 0.04$). At t4, the symptoms of the CG were reduced significantly by 4.04 points compared to t2 ($T = -4.47$, $p < 0.01$). During the same time period (t2 to t4), the EG improved 1.03 points more on this OQ-45 scale than the CG ($4.04 + 1.03 = 5.07$), although no statistical significance was reached ($T = -0.83$; $p = 0.41$). Between t2 and t5, the CG once again improved significantly by 5.10 points on the OQ-45 symptom distress scale ($T = -5.56$; $p < 0.01$). The symptom change was not significantly enhanced for the EG ($5.10 + 0.62 = 5.72$) compared to the CG between the beginning of the feedback intervention at t2 and the last available assessment point t5 ($T = -0.49$; $p = 0.62$).

On the OQ-45 interpersonal relations scale, the CG improved significantly from t1 to t2, from t2 to t4, and from t2 to t5, although not between t2 and t3. On this scale, the changes found in the EG did not differ significantly from the CG. These results are shown in Table 5.

Table 6 shows the results for the OQ-45 social performance scale. The CG improved significantly between each of the investigated assessment points (t1 → t2, t2 → t3, t2 → t4, t2 → t5) on this scale. The therapeutic process of the EG did not deviate significantly from the therapeutic process of the CG on this scale.

Discussion

In this explorative study, therapists received for the first time not only patient progress feedback for their OT patients, but also CST. This combined feedback was compared against treatment-as-usual in two psychosomatic in-patient clinics. According to the empirical algorithm of the OQ-45 [6], 82.94% (209 of 252 patients) of the psychosomatically treated in-patients were labeled as OT. This rate is in accordance with the OT rates (between 67.0% and 89.2%) found in studies from the U.S. [18].

In our study, a significant feedback effect on OT patients was found only on one of the four OQ-45 scales (OQ-45 symptom distress scale)

Table 3

Fixed effects of the level-2-model with the OQ-45 total scale as dependent variable.

Parameter	Estimate OQ-45 total scale	SE	df	t-Statistics	p-Value
t2: CG	72.65	2.64	207.68	27.54	<0.01
t2: EG vs. CG	1.58	3.62	208.15	0.44	0.66
t2 → t1: CG	7.57	1.19	202.86	6.35	<0.01
t2 → t1: EG vs. CG	-0.97	1.64	204.07	-0.06	0.95
t2 → t3: CG	-3.03	1.03	200.81	-2.93	<0.01
t2 → t3: EG vs. CG	-1.96	1.42	200.98	-1.38	0.17
t2 → t4: CG	-6.46	1.32	200.52	-4.91	<0.01
t2 → t4: EG vs. CG	-0.69	1.81	203.14	-0.38	0.70
t2 → t5: CG	-7.98	1.32	204.87	-6.05	<0.01
t2 → t5: EG vs. CG	-0.77	1.82	206.06	-0.04	0.97

Abbreviations: OQ-45 = "Outcome Questionnaire", CG = control group, EG = experimental group, SE = standard error, df = degrees of freedom, t1 = intake week, t2 = intake week + 1 week (first feedback), t3 = intake week + 2 weeks, t4 = intake week + 3 weeks, t5 = week of the last available OQ-45 score.

Table 4

Fixed effects of the level-2-model with the OQ-45 symptom distress scale as dependent variable. See text for details.

Parameter	Estimate OQ-45 symptom distress scale	SE	df	t-Statistics	p-Value
t2: CG	44.65	1.72	207.74	26.01	<0.01
t2: EG vs. CG	1.37	2.36	208.28	0.58	0.56
t2 → t1: CG	5.07	0.79	203.07	6.43	<0.01
t2 → t1: EG vs. CG	-0.41	1.09	204.26	-0.38	0.71
t2 → t3: CG	-1.56	0.70	201.07	-2.23	0.03
t2 → t3: EG vs. CG	-2.12	0.96	201.16	-2.20	0.03
t2 → t4: CG	-4.04	0.90	201.72	-4.47	<0.01
t2 → t4: EG vs. CG	-1.03	1.25	204.26	-0.83	0.41
t2 → t5: CG	-5.10	0.92	203.55	-5.56	<0.01
t2 → t5: EG vs. CG	-0.62	1.26	204.73	-0.49	0.62

Abbreviations: OQ-45 = "Outcome Questionnaire", CG = control group, EG = experimental group, SE = standard error, df = degrees of freedom, t1 = intake week, t2 = intake week + 1 week (first feedback), t3 = intake week + 2 weeks, t4 = intake week + 3 weeks, t5 = week of the last available OQ-45 score.

Table 5
Fixed effects of the level-2-model with the OQ-45 interpersonal relations scale as dependent variable.

Parameter	Estimate OQ-45 interpersonal relations scale	SE	df	t-Statistics	p-Value
t2: CG	13.82	0.68	207.71	20.25	<0.01
t2: EG vs. CG	−0.01	0.94	208.36	−0.01	0.99
t2 → t1: CG	1.65	0.39	206.43	4.28	<0.01
t2 → t1: EG vs. CG	−0.14	0.53	207.64	−0.26	0.80
t2 → t3: CG	−0.53	0.34	201.91	−1.59	0.11
t2 → t3: EG vs. CG	−0.36	0.46	202.14	−0.78	0.44
t2 → t4: CG	−1.13	0.39	201.87	−2.88	<0.01
t2 → t4: EG vs. CG	−0.21	0.54	204.62	−0.38	0.70
t2 → t5: CG	−1.57	0.36	207.56	−4.37	<0.01
t2 → t5: EG vs. CG	0.29	0.49	208.75	0.58	0.56

Abbreviations: OQ-45 = "Outcome Questionnaire", CG = control group, EG = experimental group, SE = standard error, df = degrees of freedom, t1 = intake week, t2 = intake week + 1 week (first feedback), t3 = intake week + 2 weeks, t4 = intake week + 3 weeks, t5 = week of the last available OQ-45 score.

for the duration of only one week (t2 → t3), reaching merely a very small effect size ($g = 0.12$; $M = 2.12$ points difference on this OQ-45 scale). Therefore, the clinical relevance of providing feedback to therapists in psychosomatic in-patient therapy seems less strong for OT patients than for NOT patients, whose average outcome could be enhanced substantially through the use of patient progress feedback and CST by $M = 12.50$ points ($g = 0.54$) on the OQ-45 total scale [15] in this overall project.

The lack of significant feedback effects on OT patients cannot be attributed to differences between the groups regarding demographic variables and psychological distress at time of intake. Furthermore, the EG did not differ from the CG in the amount of change on the OQ-45 scales before the feedback intervention started – from the intake week (t1) to the following week (t2). Thus, the very small feedback effect found on the OQ-45 symptom distress scale between t2 and t3 cannot be attributed to different patterns of change between the two groups during the first assessment phase (t1 to t2). As examples: If the EG had improved more than the CG from t1 to t2, the EG could for instance have been more motivated for further change. If the EG on the other hand had improved less than the CG from t1 to t2, the higher distress level at t2 could have been viewed as allowing for more changes from t2 to t3.

While our non-significant feedback effects on the outcome of OT patients are in line with two previous RCTs [10,16] and a quasi-experimental study [17], they are in contrast to a quasi-experimental study [11] and a mega-analysis [18], that showed a significantly improved outcome for OT patients when therapists received patient progress feedback. Although feedback could significantly improve the outcome of OT patients in these studies, this effect was small – as in our project – compared to the feedback effect found for NOT patients [11,18]. One should keep in mind though that only small to at the utmost moderate feedback effects should be expected for OT patients when compared to treatment-as-usual. Thus, relatively large sample sizes would be necessary to reach statistically significant results. This consideration could

contribute to explaining the ambiguous results: The sample sizes of the quasi-experimental study [11] for example exceeded 1,000 OT patients for the feedback and treatment-as-usual groups, while smaller sample sizes were analyzed in the RCTs [10,16] and in the other quasi-experimental study [17]. In our investigation, only 252 patients (209 OT patients) met the inclusion criteria although 436 patients were treated in the clinics during the duration of the study. The relatively high attrition rate (42.2%) of this study attributed to different factors (short treatment durations, high-distress severity, non-compliance) can be seen as a limitation of this study. Though in search of confounders, a drop-out analysis was performed, revealing that the in- and excluded patients did not differ from each other in the investigated variables, except for more excluded patients not giving information about their educational level. This finding hardly points to an alternative interpretation regarding confounding factors.

Another factor that could account for the different study results regarding the feedback effects on OT patients is treatment integrity [26], since the studies did not exercise control over the use of feedback. Future research is required to overcome this limitation and examine how the use of the feedback reports influences the results. In our study, OT patients whose therapists received feedback improved more on the OQ-45 symptom scale only for the one week (t2 → t3) after the first feedback reports were given to therapists (except for the 2% with missing data at t2). Therefore, one could speculate that the therapists gave more attention to the first feedback reports than to the following ones. By controlling the use of feedback, one could investigate whether the symptom improvement continues to be larger for OT patients beyond the week following the first feedback when therapists are obliged to pay attention to the feedback reports every week – if for instance the results are discussed by the whole therapeutic team. Nevertheless, it should also be kept in mind that such measures of control could also entail negative effects though demotivating the therapists.

Regarding the use of feedback, it would also be interesting for future research to examine whether the feedback effects can be enhanced by

Table 6
Fixed effects of the level-2-model with the OQ-45 social role performance scale as dependent variable.

Parameter	Estimate OQ-45 social role performance scale	SE	df	t-Statistics	p-Value
t2: CG	14.18	0.59	204.10	23.87	<0.01
t2: EG vs. CG	0.23	0.82	205.04	0.29	0.78
t2 → t1: CG	0.86	0.38	202.69	2.25	0.03
t2 → t1: EG vs. CG	0.44	0.52	203.82	0.84	0.41
t2 → t3: CG	−0.94	0.37	109.35	−2.54	0.01
t2 → t3: EG vs. CG	0.52	0.51	190.71	1.01	0.32
t2 → t4: CG	−1.29	0.38	187.72	−3.39	<0.01
t2 → t4: EG vs. CG	0.55	0.52	190.15	1.05	0.29
t2 → t5: CG	−1.32	0.40	195.85	−3.29	<0.01
t2 → t5: EG vs. CG	0.25	0.55	197.09	0.46	0.65

Abbreviations: OQ-45 = "Outcome Questionnaire", CG = control group, EG = experimental group, SE = standard error, df = degrees of freedom, t1 = intake week, t2 = intake week + 1 week (first feedback), t3 = intake week + 2 weeks, t4 = intake week + 3 weeks, t5 = week of the last available OQ-45 score.

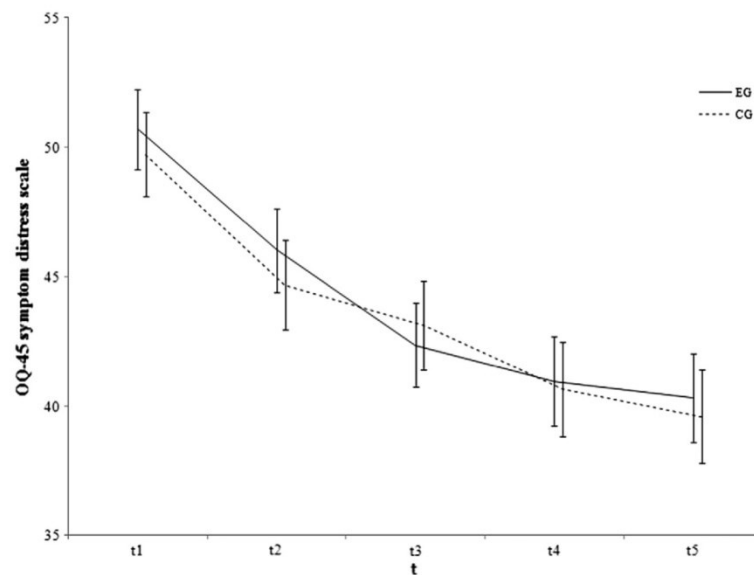


Fig. 2. Parameter estimates of the level-2-model for the symptom distress scale of the "Outcome Questionnaire" (OQ-45) at t1 (intake week), t2 (intake week + 1 week), t3 (intake week + 2 weeks), t4 (intake week + 3 weeks), and t5 (week of the last available OQ-45 score) \pm 1 standard error (SE) for the experimental group (EG) and the control group (CG).

providing feedback to all therapists involved in the multi-professional psychosomatic in-patient treatment of a given patient, and not only – as done in this study – to the therapists responsible for individual therapy. Also the effects of providing feedback not only to therapists but also to patients have not yet been scrutinized in psychosomatic therapy. In counseling, no beneficial effect of feedback could be found in such cases where feedback was also given to patients [11,12], whereas in outpatient therapy, feedback was more effective when given to both therapists and patients [27,28].

An important difference between the existing studies investigating feedback effects on OT patients and our study is that we provided not only patient progress feedback but also CST to therapists for their OT patients. Maybe adding the current CST of the OQ-Analyst to patient progress feedback is less powerful for OT patients than for NOT patients or compared to giving patient progress feedback alone. As the CST of the OQ-Analyst were initially designed for NOT patients, the CST feedback provides more information on negative than on positive results on the ASC scales (items that indicate problems are displayed and a red warning signal is given for each scale, when a patient scores below a cut-off, whereas no special signal is given for positive results). When therapists receive patient progress feedback and CST for their OT patients, the patient progress feedback is per definition positive (patient progress measured with the OQ-45 is on-track), whereas the CST can contain negative feedback. Due to the general human tendency to focus more on problems, it is conceivable that negative CST feedback could for instance also motivate the therapists to solve the problems displayed in the CST feedback (e.g. enhancing social support), instead of continuing with specific therapeutic interventions (e.g. confrontation) that had contributed to the patient's progress resulting in being on-track. This change of the therapeutic focus could also negatively influence the further progress of OT patients. To test the hypothesis that patient progress feedback alone is more helpful for OT patients than the combination of patient progress feedback and this kind of CST feedback should be investigated in future research with OT patients.

A limitation of this study is related to the applied expected recovery curves of the empirical OQ-45 algorithm. They are based upon patients from different psychotherapy settings from the U.S. Expected recovery

curves based solely on psychosomatically treated in-patients might allow assessing NOT patients and OT patients more precisely in psychosomatic in-patient treatment. There are some arguments for the generalizability of the expected recovery curves from the U.S. to German psychosomatic in-patient treatment though: The percentages of NOT patients versus OT patients were comparable to rates found in the U.S. studies [18]. Moreover, the empirical algorithm was effective at detecting patients at risk of deterioration quite early in psychosomatic in-patient treatment [15].

To summarize, the combination of patient progress feedback and CST feedback did not substantially enhance treatment-as-usual for psychosomatically treated OT in-patients in this explorative study. As this is the first study providing CST in addition to patient progress feedback to therapists for their OT patients, future studies are necessary to examine if our preliminary results can be replicated.

Conflict of interest

The authors report no financial or other relationships relevant to the subject of this article.

Acknowledgments

This research was funded by the Susa Young Gates University Professorship awarded to Michael J. Lambert.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jpsychores.2014.03.010>.

References

- [1] Lambert MJ. Outcome in psychotherapy: the past and important advances. *Psychotherapy* 2013;50:42–51.
- [2] Howard KI, Moras K, Brill PL, Martinovich Z, Lutz W. Evaluation of psychotherapy: efficacy, effectiveness, and patient progress. *Am Psychol* 1996;51:1059–64.

- [3] Lambert MJ, Kahler M, Harmon C, Shimokawa K, Burlingame G. Administration and scoring manual for the Outcome Questionnaire–45.2. Orem, UT: American Professional Credentialing Services; 2011.
- [4] Lambert MJ, Whipple JL, Bishop MJ, Vermeersch DA, Gray GV, Finch AE. Comparison of empirically derived and rationally derived methods for identifying clients at risk for treatment failure. *Clin Psychol Psychother* 2002;9:149–64.
- [5] Spielmanns G, Masters KS, Lambert MJ. A comparison of rational versus empirical methods in prediction of negative psychotherapy outcome. *Clin Psychol Psychother* 2006;13:202–14.
- [6] Finch AE, Lambert MJ, Schaalje BG. Psychotherapy quality control: the statistical generation of expected recovery curves for integration into an early warning system. *Clin Psychol Psychother* 2001;8:231–42.
- [7] Lambert MJ. Helping clinicians to use and learn from research-based systems: the OQ-Analyst. *Psychotherapy* 2012;49:109–14.
- [8] National Registry of Evidence-based Programs and Practices (NREPP). SAMHSA's national review of evidence-based practices – OQ-Analyst. OQ®-Analyst. Salt Lake City, UT: OQ Measures; 2009. [www.oqmeasures.com].
- [9] Lambert MJ, Bailey R, Kimball K, Shimokawa K, Harmon SC, Slade K. Clinical support tools manual-brief version-40. Salt Lake City: OQ Measures; 2007.
- [10] Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ. Improving the effects of psychotherapy: the use of early identification of treatment and problem-solving strategies in routine practice. *J Couns Psychol* 2003;50:59–68.
- [11] Harmon S, Lambert MJ, Smart DM, Hawkins E, Nielsen SL, Slade K, et al. Enhancing outcome for potential treatment failures: therapist–client feedback and clinical support tools. *Psychother Res* 2007;17:379–92.
- [12] Slade K, Lambert MJ, Harmon S, Smart DW, Bailey R. Improving psychotherapy outcome: the use of immediate electronic feedback and revised clinical support tools. *Clin Psychol Psychother* 2008;15:287–303.
- [13] Simon W, Lambert MJ, Harris MW, Busath G, Vazquez A. Providing patient progress information and clinical support tools to therapists: effects on patients at risk of treatment failure. *Psychother Res* 2012;22:638–47.
- [14] Crits-Christoph P, Ring-Kurtz S, Hamilton JL, Lambert MJ, Gallop R, McClure B, et al. A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *J Subst Abuse Treat* 2012;42:301–9.
- [15] Probst T, Lambert MJ, Loew TH, Dahlbender RW, Göllner R, Tritt K. Feedback on patient progress and clinical support tools for therapists: improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice. *J Psychosom Res* 2013;75:255–61.
- [16] Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins EJ. The effects of providing therapists with feedback on patient progress during psychotherapy: are outcomes enhanced? *Psychother Res* 2001;11:49–68.
- [17] Lambert MJ, Whipple JL, Vermeersch DA, Smart DW, Hawkins EJ, Nielsen SL, et al. Enhancing psychotherapy outcomes via providing feedback on client progress: a replication. *Clin Psychol Psychother* 2002;9:91–103.
- [18] Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *J Consult Clin Psychol* 2010;78:298–311.
- [19] Lambert MJ, Hannover W, Nisslmüller K, Richard M, Kordy H. Questionnaire on the results of psychotherapy: reliability and validity of the German translation of the Outcome Questionnaire 45.2 (OQ-45.2). *Z Klin Psychol Psychother* 2002;31:40–7.
- [20] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures. Report of the ISPOR task force for translation and cultural adaptation. *Value Health* 2005;8:94–104.
- [21] Singer JD, Willett JB. Applied longitudinal data analysis. Modeling change and event occurrence. London: Oxford University Press; 2003.
- [22] Göllner R, Gollwitzer M, Heider J, Zaby A, Schröder A. Analysis of longitudinal data with hierarchical linear models. *Z Klin Psychol Psychother* 2010;39:179–88.
- [23] Probst T, Loew TH, Dahlbender RW, Tritt K. Distress severity during the intake week and its weekly change in psychodynamic-oriented psychosomatic in-patient treatment. *Die Psychodynamische Psychotherapie*; 2014 [in press] <http://www.schattauer.de/de/magazine/uebersicht/zeitschriften-a-z/pdp.htm>.
- [24] Durlak JA. How to select, calculate, and interpret effect sizes. *J Pediatr Psychol* 2009;34:917–28.
- [25] World Health Organization WHO. ICD-10: the ICD-10 classification of mental and behavioural disorders. Geneva: WHO; 1992.
- [26] Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: issues and research recommendations. *Clin Psychol Sci Pract* 2005;12:365–83.
- [27] Hawkins EJ, Lambert MJ, Vermeersch DA, Slade KL, Tuttle KC. The therapeutic effects of providing patient progress information to therapists and patients. *Psychother Res* 2004;14:308–27.
- [28] De Jong K, Timman R, Hakkaart-Van Roijen L, Vermeulen P, Kooiman K, Passchier J, et al. The effect of outcome monitoring feedback to clinicians and patients in short and long-term psychotherapy: a randomized controlled trial. *Psychother Res* 2014 [Epub ahead of print].

III. 3. Publikation 3

Probst, T., Loew, T. H., Dahlbender, R. W., & Tritt, K. (2014). Belastungsgrade zur Aufnahmewoche und wöchentliche Änderungsmuster in der stationären Psychosomatik. *Die Psychodynamische Psychotherapie*, 13, 75-90.

Link: <http://www.schattauer.de/de/magazine/uebersicht/zeitschriften-a-z/pdp/inhalt/archiv/issue/special/manuscript/21247/show.html>

Belastungsgrade zur Aufnahmewoche und wöchentliche Änderungsmuster in der stationären Psychosomatik

Thomas Probst; Thomas H. Loew; Reiner W. Dahlbender; Karin Tritt

Schlüsselwörter

Stationäre Psychosomatik, Therapieprozess, Monitoring, wöchentliche Therapieverläufe

Keywords

In-patient treatment, therapeutical process, monitoring, weekly patterns of change

Zusammenfassung

Methode: Mittels zweier Fragebögen (Outcome Questionnaire „OQ-45“ und ICD-10 Symptomrating „ISR“), die die Patienten wöchentlich während ihrer psychodynamisch-orientierten stationären psychosomatischen Behandlung ausfüllten, wurden der Schweregrad behandlungsrelevanter Kriterien während der Aufnahmewoche (Woche 0) bestimmt und deren Veränderung in den folgenden drei Wochen (Woche 1, Woche 2, Woche 3) analysiert. **Ergebnisse:** Die Mehrheit der Patienten (58,1%–91,7%) schätzte sich in den Skalen beider Fragebögen in Woche 0 als belastet ein. Die Gesamtbelastung durch psychische Beschwerden, die globale Symptombelastung, die interpersonellen und sozialen Probleme verbesserten sich bereits zur Woche 1 ($p<0,05$). Die spezifische Symptombelastung veränderte sich differenziell: Die depressive Symptomatik und die Angstsymptomatik besserten sich schon zur Woche 1 ($p<0,05$). Die Somatisierungssymptomatik verbesserte sich dagegen erst zur Woche 3 ($p<0,05$). Die Zwangs- und Essstörungssymptomatik erreichte zwischen den hier untersuchten vier Messzeitpunkten noch keine signifikante Verbesserung ($p>0,05$).

Summary

Method: Patients of two psychodynamic-oriented psychosomatic in-patient clinics filled out two questionnaires (Outcome Questionnaire „OQ-45“ and ICD-10 Symptomrating „ISR“) on a weekly basis during their stay. The scores of the questionnaires were used to assess the severity of distress during the intake week (week 0) and to evaluate its change in the following three weeks (week 1, week 2, week 3). **Results:** The majority of patients (58,1%–91,7%) showed dysfunctional scores on the scales of both questionnaires at week 0. Global psychological distress, global distress due to symptoms as well as interpersonal and social problems improved already at week 1 ($p<0,05$). Specific types of symptom distress however, improved at different rates: While depressive symptoms and anxiety symptoms already improved at week 1 ($p<0,05$), symptoms typical for somatoform disorders improved first at week 3 ($p<0,05$). Symptoms of eating disorders and of obsessive-compulsive disorders did not change significantly between our four assessment points ($p>0,05$).

Distress severity during the intake week and its weekly change in psychosomatic in-patient treatment

PDP 2014; 13: 75–90

Hintergrund

Während im Rahmen der Eingangsdiagnostik geprüft wird, welche Symptome bzw. Diagnosen in welcher Ausprägung vorliegen und ob eine Indikation zur Behandlung gegeben ist, dient die Änderungsdiagnostik der Erfassung von therapeutischen Veränderungen über die Zeit. Im Bereich der Psychotherapie werden zur Eingangs- und Änderungsdiagnostik häufig psychometrische Fragebögen herangezogen, die der Patient selbst bearbeitet. Diese psychometrischen Messungen können in verschiedenen Forschungsansätzen zum Einsatz kommen: Bei der „treatment-focused research“ werden jeweils zu Beginn und Ende der Behandlung Statusmessungen erhoben, um die Effektivität einer Behandlung zu bewerten. Um Langzeiteffekte einer Behandlung zu evaluieren, finden im Rahmen dieses Ansatzes auch katamnestische Erhebungen statt. Bei der „patient-focused research“ (Howard et al. 1996) dagegen werden behandlungsrelevante Kriterien kontinuierlich während der Behandlung (z.B. jede Woche) erhoben (Monitoring). Dieses Monitoring ermöglicht eine genauere Untersuchung dessen, was sich bei welchen Patienten während der Behandlung zu welchem Zeitpunkt verändert (Kopta et al. 1994). Zudem können die Monitoringergebnisse an die Behandler rückgemeldet werden (Feedback), wodurch Behandlungen in der psychodynamisch-orientierten stationären Psychosomatik (Probst et al. 2013; in press) sowie in anderen Settings (Shimokawa et al. 2010) v.a. für Patienten mit hohem Risiko einer Verschlechterung optimiert werden können.

In dieser Studie wurden zwei unterschiedliche Fragestellungen untersucht: Erstens, wie belastet sich Patienten während der Aufnahmewoche einer psychodynamisch-orientierten stationären psychosomatischen Behandlung bzgl. verschiedener behandlungsrelevanter Kriterien (globale und spezifische psychische Symptome, interpersonelle und soziale Probleme) einschätzen (Eingangsdiagnostik). Zweitens wird anhand von wöchentlichen Monitoringdaten analysiert, in welcher Form sich die behandlungsrelevanten

Kriterien zu Beginn der Behandlung – zwischen Aufnahmewoche und den folgenden drei Behandlungswochen – verändern (Änderungsdiagnostik). In früheren Studien bzgl. der Effektivität stationärer psychosomatischer Behandlung (vgl. Liebherz, Rabung 2013; Steffanowski et al. 2007) wurden v.a. Veränderungen zwischen Aufnahme und Entlassung untersucht. Teilweise wurden in früheren Studien auch Zwischenmessungen während der Behandlung (z.B. vier Wochen nach Aufnahme) und Katamnesen erhoben. Melchior (2011) untersuchte das wöchentliche Verlaufsmuster in stationären psychosomatischen Kliniken hinsichtlich globaler psychischer Symptombelastung, um Verlaufsgruppen zu bestimmen. Im Gegensatz dazu wird in dieser Untersuchung ein erster Versuch unternommen, wöchentliche Verlaufsmuster, die sich während der ersten Wochen psychodynamisch-orientierter stationärer psychosomatischer Behandlung ereignen, hinsichtlich differenzieller Veränderungen genauer zu untersuchen. Das Phasenmodell therapeutischer Veränderungen (Howard et al. 1993) würde erwarten lassen, dass sich hinsichtlich der behandlungsrelevanten Kriterien zuerst Symptome und danach interpersonelle und soziale Probleme verbessern. Haase et al. (2008) haben in diesem Zusammenhang bei psychodynamisch-orientierter stationärer psychosomatischer Behandlung gezeigt, dass sich zur Zwischenmessung vier Woche nach Aufnahme besonders psychische Symptome und auch einige interpersonellen Probleme verbesserten. Wir wollten in dieser Arbeit im Rahmen der Fragestellung zur Änderungsdiagnostik genauer untersuchen, wann sich im wöchentlichen Verlauf welche psychischen Symptombelastungen verändern und ob sich interpersonelle sowie soziale Probleme bereits in den ersten Wochen der Behandlung verbessern und wann dies geschieht.

Methode

Im Rahmen der Darstellung der Methodik dieser Studie wird zuerst das Studiendesign erläutert,

gefolgt von einer Beschreibung der zwei eingesetzten Fragebögen und einer Darstellung der verwendeten statistischen Verfahren. Zum Abschluss folgt eine Beschreibung der Stichproben dieser Studie.

Studiendesign

In zwei psychodynamisch-orientierten psychosomatischen Kliniken (Psychosomatische Abteilung der Klinik Donaustauf und „Klinik am schönen Moos“ Bad Saulgau) bearbeiteten Patienten wöchentlich u.a. den „Outcome Questionnaire“ (OQ-45; Lambert et al. 2011) und das ICD-10 Symptomrating (ISR; Tritt et al. 2008). In der Klinik Donaustauf nahmen alle Patienten der Psychosomatikstation an der Studie teil, die vom 01.10.2010 bis zum 06.07.2012 stattfand. Vom 16.01.2012 bis zum 06.04.2012 beteiligten sich in Bad Saulgau Patienten mit privater Krankenversicherung an der Studie. Die Fragebögen wurden den Patienten während der Behandlung einmal pro Woche freitags ausgehändigt mit der Bitte, sie bis Montag auszufüllen. Die Daten des OQ-45 wurden montags in die Monitoring- und Feedbacksoftware OQ-Analyst (Lambert, 2012) übertragen und für die Patienten, die bei der Aufnahme zufällig der Feedbackgruppe zugeordnet wurden, bekamen die behandelnden Einzeltherapeuten die Feedbackberichte des OQ-Analyst in einem geschlossenen Kuvert ausgehändigt. Die Feedbackeffekte dieses Projekts wurden in anderen Originalarbeiten dargestellt (Probst et al. 2013; Probst et al., in press). Das Behandlungssetting in den beteiligten Kliniken beinhaltete Einzel- und Gruppenpsychotherapien, Entspannungs- und Achtsamkeitstrainings, Körper-, Musik-, Kunst-, Tanz-, Physiotherapien und wenn nötig, medikamentöse Behandlungen (multimodale Komplexbehandlung). Insgesamt verbrachte jeder Patient pro Behandlungswoche mehr als 24 Stunden in verschiedenen therapeutischen Behandlungen.

Fragebögen

Folgende zwei Fragebogen wurden in dieser Studie eingesetzt:

Outcome Questionnaire (OQ-45): Der Selbsteinschätzungsfragebogen OQ-45 (Lambert et al. 2011) besteht aus 45 Items, die alle auf einer fünfstufigen Likert-Skala vom Patienten bewertet werden. Alle Items beziehen sich auf den Zeitraum der letzten Woche. Aus allen 45 Items wird die „Gesamtskala“ berechnet, 25 Items entfallen auf die Skala „Symptombelastung“, 11 Items gehören zur Skala „Zwischenmenschliche Beziehungen“ und 9 Items werden in die Skala „Soziale Integration“ verrechnet. Höhere Werte stehen für eine höhere Beeinträchtigung. Die psychometrischen Qualitäten der deutschen Version des OQ-45 berichteten Lambert et al. (2002). U.a. wurden folgende Konsistenzkoeffizienten (Alpha) und Retestreliabilitäten (r) gefunden: „Gesamtskala“ (Alpha=0,93; r=0,88), Skala „Symptombelastung“ (Alpha=0,90; r=0,87), Skala „Zwischenmenschliche Beziehungen“ (Alpha=0,81; r=0,81) und Skala „Soziale Integration“ (Alpha=0,59; r=0,71). Die „Gesamtskala“ wurde in dieser Studie zur Messung der globalen Gesamtbelastung durch verschiedene psychische Probleme verwendet. Die Skala „Symptombelastung“ diente zur Operationalisierung der globalen Belastung durch psychische Symptome. Mit der Skala „Zwischenmenschliche Beziehungen“ sollten zwischenmenschliche Probleme und mit der Skala „Soziale Integration“ soziale Probleme erhoben werden.

ICD-10 Symptomrating (ISR): Das ISR (Tritt et al. 2008; unter www.iqp-online.de im Downloadbereich lizenzfrei verfügbar) ist ein Selbsteinschätzungsfragebogen, der 29 Items umfasst und eng auf dem Kapitel V des ICD-10 basiert. Alle Items werden bzgl. der letzten zwei Wochen auf einer fünfstufigen Likert-Skala geratet. Die Items verteilen sich wie folgt auf die ISR-Syndromskalen: 4 Items für die „Depressionsskala“, 4 Items für die „Angstskala“, 3 Items für die „Zwangsskala“, 3 Items für die „Somatisierungsskala“ und 3 Items für die „Essstörungsskala“. Für jede

Syndromskala wird ein arithmetischer Mittelwert bestimmt. Zudem gibt es 12 Zusatzitems, die keiner Skala zugeordnet sind, da sie entweder bei verschiedenen psychischen Syndromen vorkommen können oder sie erste Hinweise auf das Vorliegen spezifischer Syndrome liefern sollen. Aus allen Syndromskalen sowie den 12 Zusatzitems wird eine „Gesamtskala“ errechnet, die Hinweise auf die globale subjektive symptomatische Belastung des Patienten geben soll. Höhere Werte stehen im ISR für eine höhere Belastung. Ergebnisse zu den psychometrischen Qualitäten des ISR berichteten u.a. Fischer et al. (2010; 2011) und Tritt et al. (2010a). Die Retestreliabilitäten der einzelnen Skalen variierten von $r=0,70$ bis $r=0,94$ (Fischer et al. 2011). Mit der „Gesamtskala“ des ISR sollte in dieser Studie – ebenso wie mit der OQ-45 Skala „Symptombelastung“ – die globale Belastung durch psychische Symptome gemessen werden. Die ISR-Syndromskalen wurden zur Operationalisierung von spezifischen Symptombelastungen verwendet.

Statistische Auswertung

Die statistische Auswertung wurde mit SPSS 21 durchgeführt. Alle durchgeführten statistischen Tests erfolgten zweiseitig mit einem Signifikanzniveau von 5%. Für die Stichprobenbeschreibungen wurden Häufigkeiten (n, %), Mittelwerte (M) und Standardabweichungen (SD) berechnet. Die bei der Eingangs- und Änderungsdiagnostik eingesetzten statistischen Verfahren werden im Folgenden dargestellt.

Eingangsdiagnostik

Um den Schweregrad der Patienten während der Aufnahmewoche (Woche 0) zu kategorisieren, wurden die Schwellenwerte für die Skalen des OQ-45 aus Lambert et al. (2002) verwendet. Dies sind für die einzelnen OQ-45-Skalen:

- OQ-45 „Gesamtskala“: Ab 59 Punkten gilt ein Wert als auffällig
- OQ-45 Skala „Symptombelastung“: Punkte ab 33 gelten als auffällig

- OQ-45 Skala „Zwischenmenschliche Beziehungen“: Mind. 13 Punkte sind nötig, um als auffällig klassifiziert zu werden
- OQ-45 Skala „Soziale Integration“: Ein auffälliger Wert liegt ab 12 Punkten vor

Für das ISR wurden die Schwellenwerte verwendet, wie sie im Rahmen der ISR-Normierung festgelegt wurden (Tritt et al. 2010b).

- ISR „Gesamtskala“: Werte ab 0,5 weisen auf eine psychische Störung hin
- ISR „Depressionsskala“: Werte ab 0,75 gelten als Hinweis auf das Vorliegen einer depressiven Störung
- ISR „Angstskala“: Werte ab 0,75 deuten auf eine Angststörung hin
- ISR „Somatisierungsskala“: Werte ab 0,33 werden als Hinweise für eine somatoforme Störung gewertet
- ISR „Essstörungsskala“: Bei Werten ab 0,33 liegen Hinweise für eine Essstörung vor

Änderungsdiagnostik

Die OQ-45 und ISR-Daten der Aufnahmewoche und der folgenden drei Behandlungswochen wurden mittels Varianzanalysen mit Messwiederholung analysiert, dabei war der Messwiederholungsfaktor „Behandlungsverlauf“ 4-fach gestuft (Woche 0, Woche 1, Woche 2, Woche 3). Die korrigierten Greenhouse-Geisser-Werte wurden herangezogen, sofern die Sphärizitätsannahme verletzt wurde (signifikanter Mauchly-Test). Um zu analysieren, wie sich die Patienten zwischen den einzelnen Wochen in den Fragebogenskalen veränderten, wurden im Rahmen der Varianzanalysen paarweise Vergleiche mit Bonferroni-Korrektur verwendet. Zudem wurden Effektstärken (d) nach folgender Formel berechnet, um die Veränderung zwischen den einzelnen Wochen zu quantifizieren:

$$d = (M_{\text{WocheX}} - M_{\text{WocheX+1}}) / SD_{\text{WocheX}}$$

In die statistischen Analysen zur Änderungsdiagnostik wurden nur die Patienten eingeschlossen, für die zu allen vier Messzeitpunkten (Woche 0, Woche 1, Woche 2, Woche 3) OQ-45- und ISR-Werte vorhanden waren. Um die Vergleichbarkeit zwischen den zur Ände-

rungsdiagnostik ein- und ausgeschlossenen Patienten zu bewerten, wurde eine Drop-Out-Analyse durchgeführt. Bei dieser Drop-Out-Analyse wurden für nominale Variablen Exakte Tests nach Fisher (FET) berechnet und für metrische Variablen t-Tests für unabhängige Stichproben. Bei den t-Tests für unabhängige Stichproben wurden die korrigierten Werte verwendet, wenn keine Varianzgleichheit vorlag (signifikanter Levene-Test).

Stichprobe

In dieser Studie wurden zwei Stichproben untersucht, die im Folgenden beschrieben werden: Die erste Stichprobe besteht aus Patienten mit vorhandener Messung zur Aufnahmewoche (Stichprobe zur Eingangsdiagnostik). Bei der zweiten Stichprobe handelt es sich um eine Substichprobe der ersten Stichprobe, bestehend aus den Patienten, für die Messungen sowohl zur Aufnahmewoche als auch zu den folgenden drei Behandlungswochen vorhanden waren (Stichprobe zur Änderungsdiagnostik). Während zur Aufnahmewoche noch Daten für 277 Patienten vorlagen, waren für 127 Patienten (46% von 277) verwertbare wöchentliche OQ-45 und ISR-Daten zu allen vier Messzeitpunkten vorhanden. Dieser Stichprobenschwund hatte zwei Gründe: Die eine Hälfte der ausgeschlossenen 150 Patienten war – aufgrund einer zeitlich eng begrenzten Krisenintervention – nicht so lange in stationärer Behandlung, wie es nötig gewesen wäre, um zu allen vier Messzeitpunkten vorhandene Datensätze liefern zu können. Die andere Hälfte füllte nicht zu allen drei weiteren Messzeitpunkten (sondern nur zu einem oder zwei) die beiden Fragebögen an den Wochenenden verwertbar aus (z.B. aufgrund von Belastungsproben an den Wochenenden). Bei der Darstellung der Stichprobe zur Änderungsdiagnostik werden auch die Ergebnisse der Drop-Out-Analyse dargestellt.

Eingangsdiagnostik

Tabelle 1 gibt einen Überblick über die sozio-demografischen und diagnostischen Merkmale

der Stichprobe zur Eingangsdiagnostik. Bei den Diagnosen wurden jeweils alle Diagnosen (nicht nur die Erstdiagnosen) berücksichtigt, um den Komorbiditäten Rechnung zu tragen. Die Diagnosen wurden nach Kapitel V des ICD-10 von den Klinikteams gestellt.

Änderungsdiagnostik

Über die sozio-demografischen und diagnostischen Merkmale der Stichprobe zur Änderungsdiagnostik sowie über die Drop-Out Analyse informiert Tabelle 2.

Ergebnisse

Nach einer Darstellung des Schweregrads der subjektiven Belastungen der Patienten während der Aufnahmewoche (Eingangsdiagnostik), werden die Veränderungen zwischen Aufnahmewoche und den folgenden drei Wochen dargestellt (Änderungsdiagnostik).

Eingangsdiagnostik

Die globale Belastung durch psychische Beschwerden (OQ-45 Skalen und ISR „Gesamtskala“) der Stichprobe zur Eingangsdiagnostik wird in Tabelle 3 dargestellt.

Die Ergebnisse bzgl. des Ausmaßes spezifischer Symptombelastungen (ISR-Syndromskalen) während der Aufnahmewoche präsentiert Tabelle 4.

Änderungsdiagnostik

Zur Sichtung der Veränderungen der globalen Kriterien zwischen Aufnahmewoche und den folgenden drei Wochen werden in Tabelle 5 die Mittelwerte, Standardabweichungen und die Ergebnisse der Varianzanalysen mit Messwiederholung sowie die paarweisen Vergleiche für die OQ-45 Skalen und die ISR „Gesamtskala“ präsentiert.

Tabelle 6 stellt die Veränderungen hinsichtlich der spezifischen Kriterien (ISR-Syndromskalen) für die untersuchten vier Messzeitpunkte dar sowie die dazugehörigen Ergebnisse der Va-

Variable		Stichprobe zur Eingangsdiagnostik N=277
Geschlecht n (%)	weiblich	167 (60,3%)
	männlich	110 (39,7%)
Alter bei Aufnahme M (SD)		48,26 (13,66)
Schulabschluss n (%)	Hauptschule	119 (43,0%)
	Realschule	79 (28,5%)
	Abitur	57 (20,6%)
	Sonstiges	22 (7,9%)
Kumulierte ICD-10 F-Diagnosen n (%)	Depressive Störung F32 / F33 / F34 / F38 / F39	179 (64,6%)
	Angststörung F40 / F41	63 (22,7%)
	Zwangsstörung F42	10 (3,6%)
	Belastungsstörung F43	35 (12,6%)
	Somatoforme Störung F45	157 (56,7%)
	Essstörung F50	39 (14,1%)
	Persönlichkeitsstörung F6	21 (7,6%)
	sonstige F-Diagnosen	60 (21,7%)
Mittlere Anzahl von ICD-10 F-Diagnosen pro Patient M (SD)		2,18 (1,01)

Tab. 1

Sozio-demographische und diagnostische Merkmale der Stichprobe zur Eingangsdiagnostik.

riananalysen mit Messwiederholung und der paarweisen Vergleiche.

In Tabelle 7 werden die Effektstärken berichtet, wie sie sich zwischen den vier untersuchten Messzeitpunkten für die OQ-45 und ISR-Skalen ergaben.

Diskussion

Diese Arbeit untersuchte zwei Fragestellungen: Erstens, wie belastet sich Patienten in psychodynamisch-orientierter stationärer psychosomati-

scher Behandlung während der Aufnahmewoche (Woche 0) hinsichtlich verschiedener behandlungsrelevanter globaler bzw. spezifischer Kriterien einschätzten. Zweitens wurde anhand von wöchentlichen Messungen versucht, differenzielle Verlaufsmuster globaler und spezifischer Kriterien zwischen Aufnahmewoche (Woche 0) und folgenden drei Behandlungswochen (Woche 1, Woche 2, Woche 3) abzubilden. Globale Kriterien waren psychische Gesamtbelastung, globale Belastung durch psychische Symptome, interpersonelle und soziale Probleme. Spezifische Kriterien waren Symptombelastun-

Tab. 2 Sozio-demographische und diagnostische Merkmale der Stichprobe zur Änderungsdiagnostik.

Variable		Stichprobe zur Änderungsdiagnostik N=126	Drop-Outs N=151	Statistik
Geschlecht n (%)	weiblich	76 (60,3%)	91 (60,3%)	p=1,00 (FET)
	männlich	50 (39,7%)	60 (39,7%)	
Alter bei Aufnahme M (SD)		46,54 (12,25)	49,70 (14,63)	t(274,995)=1,955; p=0,052
Schulabschluss n (%)	Hauptschule	53 (42,1%)	66 (44,3%)	p=0,88 (FET)
	Realschule	39 (31,0%)	40 (26,8%)	
	Abitur	26 (20,6%)	31 (20,8%)	
	Sonstiges	8 (6,3%)	12 (8,1%)	
Kumulierte ICD-10 F-Diagnosen n (%)	Depressive Störung F32 / F33 / F34 / F38 / F39	89 (70,6%)	90 (59,6%)	p=0,06 (FET)
	Angststörung F40 / F41	33 (26,2%)	30 (19,9%)	p=0,25 (FET)
	Zwangsstörung F42	7 (5,6%)	3 (2,0%)	p=0,19 (FET)
	Belastungsstörung F43	18 (14,3%)	17 (11,3%)	p=0,47 (FET)
	Somatoforme Störung F45	72 (57,1%)	85 (56,3%)	p=0,90 (FET)
	Essstörung F50	20 (15,9%)	19 (12,6%)	p=0,49 (FET)
	Persönlichkeitsstörung F6	9 (7,1%)	12 (7,9%)	p=0,82 (FET)
	sonstige F-Diagnosen	20 (15,9%)	40 (26,5%)	p=0,04 (FET)
Mittlere Anzahl von ICD-10 F-Diagnosen pro Patient M (SD)		2,25 (1,02)	2,12 (1,01)	t(275)=-1,103; p=0,271

gen spezifischer psychischer Störungen (Depressive Störungen, Angststörungen, Zwangsstörungen, somatoforme Störungen, Essstörungen).

Die Ergebnisse der Eingangsdiagnostik zeigten, dass die Patienten während der Aufnahmewoche im Mittel ähnlich hohe Scores in den OQ-45-Skalen „Symptombelastung“ und „Sozia-

le Integration“ aufwiesen wie die von Bauer (2004) untersuchten stationär psychosomatisch behandelten Patienten, jedoch etwas geringere Werte auf der OQ-45-Skala „Zwischenmenschliche Beziehungen“. Die Gesamtbelastung durch psychische Beschwerden (OQ-45 „Gesamtskala“) während der Aufnahmewoche fiel augen-

Fragebogen – Skala	Belastung	
OQ-45 „Gesamtskala“	M (SD)	78,26 (25,87)
	unauffällig n (%)	65 (23,5%)
	auffällig n (%)	212 (76,5%)
OQ-45 Skala „Symptombelastung“	M (SD)	48,96 (16,52)
	unauffällig n (%)	50 (18,1%)
	auffällig n (%)	227 (81,9%)
OQ-45 Skala „Zwischenmenschliche Beziehungen“	M (SD)	14,56 (7,03)
	unauffällig n (%)	108 (39,0%)
	auffällig n (%)	169 (61,0%)
OQ-45 Skala „Soziale Integration“	M (SD)	14,74 (5,88)
	unauffällig n (%)	88 (31,8%)
	auffällig n (%)	189 (68,2%)
ISR „Gesamtskala“	M (SD)	1,39 (0,70)
	keine Hinweise auf psychische Störungen n (%)	27 (9,7%)
	Hinweise auf psychische Störungen n (%)	250 (90,3%)

Tab. 3

Globale Belastung durch psychische Beschwerden der Stichprobe zur Eingangsdiagnostik in den OQ-45-Skalen und der ISR „Gesamtskala“.

scheinlich bei unserer Stichprobe etwas geringer ($M=78,3$ vs. $M=81,8$) aus als bei Bauer (2004) und etwas höher ($M=78,3$ vs. $M=74,2$) als bei Percevic et al. (2006). Verglichen mit Patienten in stationär psychiatrischer Behandlung (Puschner et al. 2009) zeigten die Patienten unserer Studie vergleichbare OQ-45-Gesamtscores während der Aufnahmewoche ($M=78,3$ vs. $M=81,4$ bzw. $M=76,6$). Im Vergleich zu anderen Studien zum ISR (Brandt, 2010; Seidel, 2013) war der Mittelwert unserer Patienten etwas höher in der ISR „Somatisierungsskala“ ($M=1,10$ vs. $M=0,68$ bzw. $M=1,03$) und der ISR „Essstörungsskala“ ($M=0,94$ vs. $M=0,74$ bzw. $M=0,91$). Im Mittel etwas geringer scorten die Patienten unserer Studie in der ISR „Zwangsskala“ ($M=1,24$ vs. $M=1,36$

bzw. $M=1,63$) sowie in der ISR „Depressionsskala“ ($M=2,20$ vs. $M=2,42$ bzw. $M=2,59$). In der ISR „Angstsskala“ sowie in der ISR „Gesamtskala“ lagen unsere Mittelwerte im Range der beiden zitierten Studien. Hinsichtlich der Klassifikation in auffällige und unauffällige Patienten anhand der Fragebogenwerte schätzten sich die meisten Patienten bei den Skalen zur globalen psychischen Symptombelastung während der Aufnahmewoche als belastet ein: 81,9% erreichten auffällige Werte in der OQ-45 Skala „Symptombelastung“ und 90,3% in der ISR „Gesamtskala“. Am wenigsten belastet schätzten sich die Patienten hinsichtlich globaler Kriterien in der OQ-45-Skala „Zwischenmenschliche Beziehungen“ ein, in der 39,0% der Patienten unauffällige Wer-

Tab. 4

Spezifische Symptombelastung der Stichprobe zur Eingangsdiagnostik in den ISR-Syndromskalen.

Fragebogen – Skala	Belastung	
ISR „Depressionsskala“	M (SD)	2,20 (1,04)
	keine Hinweise auf depressive Störungen n (%)	23 (8,3%)
	Hinweise auf depressive Störungen n (%)	254 (91,7%)
ISR „Angstskala“	M (SD)	1,83 (1,27)
	keine Hinweise auf Angststörungen n (%)	62 (22,4%)
	Hinweise auf Angststörungen n (%)	215 (77,6%)
ISR „Zwangsskala“	M (SD)	1,24 (1,13)
	keine Hinweise auf Zwangsstörungen n (%)	100 (36,1%)
	Hinweise auf Zwangsstörungen n (%)	177 (63,9%)
ISR „Somatisierungsskala“	M (SD)	1,10 (1,12)
	keine Hinweise auf somatoforme Störungen n (%)	79 (28,5%)
	Hinweise auf somatoforme Störungen n (%)	198 (71,5%)
ISR „Essstörungsskala“	M (SD)	0,94 (1,16)
	keine Hinweise auf Essstörungen n (%)	116 (41,9%)
	Hinweise auf Essstörungen n (%)	161 (58,1%)

te erreichten. Die vergleichsweise geringere Belastung durch interpersonelle Probleme könnte damit zusammenhängen, dass relativ wenige Patienten der Stichprobe mit Persönlichkeitsstörungen (7,6%) diagnostiziert wurden, bei denen man insbesondere interpersonelle Probleme erwarten würde. Diese Ergebnisse unserer Studie stehen in Einklang mit den Ergebnissen von Bauer (2004), in deren Studie sich Patienten in stationär psychosomatischer Behandlung zu Therapie-

beginn ebenfalls am stärksten belastet in der OQ-45-Skala „Symptombelastung“ (79,4% auffällig) und am wenigsten belastet in der OQ-45-Skala „Zwischenmenschliche Beziehungen“ (34,9% funktional) einschätzten. Bzgl. der spezifischen Kriterien (ISR-Syndromskalen) gaben die Patienten während der Aufnahmewoche auf der ISR „Depressionsskala“ die höchste Belastung an: 91,7% hatten auffällige Werte. Am wenigsten belastet zeigten sich die Patienten auf der ISR

Tab. 5 Veränderungen der Stichprobe zur Änderungsdiagnostik hinsichtlich globaler Belastung durch psychische Beschwerden in den OQ-45-Skalen und der ISR „Gesamtskala“.

Fragebogen – Skala	Woche 0 M (SD)	Woche 1 M (SD)	Woche 2 M (SD)	Woche 3 M (SD)	Ergebnisse der Varianzanalysen mit Mess- wiederholung und der paarweisen Vergleiche
OQ-45 „Gesamtskala“	84,06 (26,17)	77,64 (27,69)	75,77 (29,47)	73,13 (31,70)	F(2,14, 267,26)=26,49, p<0,05 Woche 0 vs. Woche 1: p<0,05 Woche 0 vs. Woche 2: p<0,05 Woche 0 vs. Woche 3: p<0,05 Woche 1 vs. Woche 2: p=0,463 Woche 1 vs. Woche 3: p<0,05 Woche 2 vs. Woche 3: p=0,054
OQ-45 Skala „Symptom- belastung“	52,25 (16,38)	48,17 (18,17)	46,75 (18,74)	45,08 (20,45)	F(2,23, 278,80)=27,93, p<0,05 Woche 0 vs. Woche 1: p<0,05 Woche 0 vs. Woche 2: p<0,05 Woche 0 vs. Woche 3: p<0,05 Woche 1 vs. Woche 2: p=0,228 Woche 1 vs. Woche 3: p<0,05 Woche 2 vs. Woche 3: p=0,075
OQ-45 Skala „Zwischen- menschliche Beziehungen“	15,79 (7,38)	14,43 (6,98)	14,29 (7,57)	13,83 (8,08)	F(2,39, 299,18)=8,64, p<0,05 Woche 0 vs. Woche 1: p<0,05 Woche 0 vs. Woche 2: p<0,05 Woche 0 vs. Woche 3: p<0,05 Woche 1 vs. Woche 2: p=1,000 Woche 1 vs. Woche 3: p=0,650 Woche 2 vs. Woche 3: p=1,000

Tab. 5 Fortsetzung

Fragebogen – Skala	Woche 0 M (SD)	Woche 1 M (SD)	Woche 2 M (SD)	Woche 3 M (SD)	Ergebnisse der Varianzanalysen mit Mess- wiederholung und der paarweisen Vergleiche
OQ-45 Skala „Soziale Integration“	16,02 (5,81)	15,05 (6,00)	14,72 (6,49)	14,22 (5,95)	$F(2,72, 339,58)=10,17$, $p<0,05$ Woche 0 vs. Woche 1: $p<0,05$ Woche 0 vs. Woche 2: $p<0,05$ Woche 0 vs. Woche 3: $p<0,05$ Woche 1 vs. Woche 2: $p=1,000$ Woche 1 vs. Woche 3: $p=0,058$ Woche 2 vs. Woche 3: $p=0,459$
ISR „Gesamtskala“	1,50 (0,68)	1,31 (0,69)	1,31 (0,73)	1,26 (0,75)	$F(2,13, 266,67)=21,47$, $p<0,05$ Woche 0 vs. Woche 1: $p<0,05$ Woche 0 vs. Woche 2: $p<0,05$ Woche 0 vs. Woche 3: $p<0,05$ Woche 1 vs. Woche 2: $p=1,000$ Woche 1 vs. Woche 3: $p=0,315$ Woche 2 vs. Woche 3: $p=0,149$

„Essstörungsskala“. Dort erreichten 41,9% unauffällige Werte, gefolgt von der ISR „Zwangsskala“, auf der 36,1% der Patienten in den unauffälligen Bereich fielen. Dies könnte u.a. dadurch bedingt sein, dass bei relativ wenigen Patienten der Stichprobe Ess- (14,1%) bzw. Zwangsstörungen (3,6%) diagnostiziert wurden und man eigentlich nur von den entsprechend diagnostizierten Patienten auffällige Werte in diesen Skalen erwarten kann. Diese relativ geringen Prävalenzraten für Zwangsstörungen wurden auch in der Allgemeinbevölkerung (Wittchen & Jacobi, 2012) sowie in klinischen Stichproben, z.B. von Heymann et al. (2003), beobachtet. Als weiterer Faktor sollte in diesem Zusammenhang ein Resultat der ISR-Normierungsstudie angeführt wer-

den (Tritt et al. 2010b): Knapp 17% der Patienten mit einer Essstörungsdiagnose gaben dort keine Belastung in der ISR „Essstörungsskala“ an. Dieser Befund dürfte mit der klinischen Beobachtung von Krankheits-Verleugnungstendenzen dieser Patienten zusammenhängen. Schließlich könnte der Anteil von Patienten mit unauffälligen Werten generell auch dadurch beeinflusst worden sein, dass in dieser Studie zwischen 0 und 5 Behandlungstage – je nachdem an welchem Wochentag der Patient aufgenommen wurde – zwischen Aufnahmetag und erstem Messzeitpunkt (Freitag bis Sonntag nach Aufnahme) lagen: Einige Patienten könnten sich möglicherweise bereits in den Tagen zwischen Aufnahmetag und erster Messung soweit stabilisiert haben, dass sie

Tab. 6 Veränderungen der Stichprobe zur Änderungsdiagnostik hinsichtlich spezifischer Symptombelastung in den ISR-Syndromkalen.

Fragebogen – Skala	Woche 0 M (SD)	Woche 1 M (SD)	Woche 2 M (SD)	Woche 3 M (SD)	Ergebnisse der Varianzanalysen mit Messwiederholung und der paarweisen Vergleiche
ISR „Depressionskala“	2,38 (1,04)	1,96 (1,13)	1,93 (1,16)	1,80 (1,20)	F(2,43, 303,07)=29,83, p<0,05 Woche 0 vs. Woche 1: p<0,05 Woche 0 vs. Woche 2: p<0,05 Woche 0 vs. Woche 3: p<0,05 Woche 1 vs. Woche 2: p=1,000 Woche 1 vs. Woche 3: p=0,072 Woche 2 vs. Woche 3: p=0,072
ISR „Angstskala“	2,07 (1,21)	1,69 (1,27)	1,76 (1,25)	1,69 (1,21)	F(2,58, 322,99)=13,24, p<0,05 Woche 0 vs. Woche 1: p<0,05 Woche 0 vs. Woche 2: p<0,05 Woche 0 vs. Woche 3: p<0,05 Woche 1 vs. Woche 2: p=1,000 Woche 1 vs. Woche 3: p=1,000 Woche 2 vs. Woche 3: p=1,000
ISR „Zwangsskala“	1,45 (1,23)	1,33 (1,20)	1,31 (1,15)	1,32 (1,16)	F(2,62, 327,52)=1,99, p=0,124 Woche 0 vs. Woche 1: p=0,587 Woche 0 vs. Woche 2: p=0,375 Woche 0 vs. Woche 3: p=0,446 Woche 1 vs. Woche 2: p=1,000 Woche 1 vs. Woche 3: p=1,000 Woche 2 vs. Woche 3: p=1,000

Tab. 6 Fortsetzung

Fragebogen – Skala	Woche 0 M (SD)	Woche 1 M (SD)	Woche 2 M (SD)	Woche 3 M (SD)	Ergebnisse der Varianz- analysen mit Mess- wiederholung und der paarweisen Vergleiche
ISR „Somatisierungs- skala“	1,01 (1,08)	0,92 (0,99)	0,85 (1,02)	0,75 (0,97)	F(2,69, 335,89)=7,16, p<0,05 Woche 0 vs. Woche 1: p=0,776 Woche 0 vs. Woche 2: p=0,054 Woche 0 vs. Woche 3: p<0,05 Woche 1 vs. Woche 2: p=1,000 Woche 1 vs. Woche 3: p<0,05 Woche 2 vs. Woche 3: p=0,223
ISR „Essstörungen- skala“	1,04 (1,23)	0,94 (1,18)	0,98 (1,24)	1,03 (1,31)	F(2,18, 272,63)=2,01, p=0,112 Woche 0 vs. Woche 1: p=0,219 Woche 0 vs. Woche 2: p=1,000 Woche 0 vs. Woche 3: p=1,000 Woche 1 vs. Woche 2: p=1,000 Woche 1 vs. Woche 3: p=0,226 Woche 2 vs. Woche 3: p=0,687

sich schon zur ersten Messung als unauffällig einschätzten. Im Rahmen dieser Studie war es aus organisatorischen Gründen nicht möglich, bei den Patienten direkt am Aufnahmetag eine Erhebung durchzuführen und dann stets am gleichen Tag in wöchentlichen Abständen die Fragebögen erneut zu bearbeiten: Aus ökonomischen Gründen wurden die Fragebögen an einem festen Wochentag (Freitag) an die Patienten verteilt.

Insgesamt wird deutlich, dass die Patienten in der Regel ein beachtliches Ausmaß an Beeinträchtigungen während der Aufnahmewoche aufwiesen, das vergleichbar mit den Befunden anderer Studien in der Psychosomatik und Psychiatrie ausfiel.

Im Rahmen der Änderungsdiagnostik wurden die wöchentlichen Verlaufsmuster globaler und

spezifischer behandlungsrelevanter Kriterien innerhalb der ersten Behandlungswochen genauer untersucht. Zuerst werden die Ergebnisse globaler Kriterien und danach die der spezifischen Kriterien diskutiert.

Die Ergebnisse zu den globalen Kriterien zeigten, dass sich die Patienten sowohl hinsichtlich der psychischen Gesamtbelastung (OQ-45 „Gesamtskala“), als auch hinsichtlich globaler psychischer Symptome (OQ-45-Skala „Symptombelastung“ und ISR „Gesamtskala“), interpersoneller (OQ-45-Skala „Zwischenmenschliche Beziehungen“) sowie sozialer Probleme (OQ-45-Skala „Soziale Integration“) während des Erhebungszeitraums verbesserten. Von Woche 0 bis Woche 3 haben sich alle erfassten glo-

Tab. 7 Effektstärken (d) für die Veränderungen zwischen Aufnahmewoche (Woche 0) und den folgenden drei Behandlungswochen (Woche 1, Woche 2, Woche 3) für die OQ-45 und ISR Skalen.

	Fragebogen – Skala	Woche 1 d	Woche 2 d	Woche 3 d
Woche 0	OQ-45 „Gesamtskala“	0,25	0,32	0,42
	OQ-45 Skala „Symptombelastung“	0,25	0,34	0,44
	OQ-45 Skala „Zwischenmenschliche Beziehungen“	0,18	0,20	0,27
	OQ-45 Skala „Soziale Integration“	0,17	0,22	0,31
	ISR „Gesamtskala“	0,27	0,28	0,36
	ISR „Depressionsskala“	0,40	0,43	0,55
	ISR „Angstskala“	0,31	0,28	0,31
	ISR „Zwangsskala“	0,10	0,12	0,11
	ISR „Somatisierungsskala“	0,09	0,15	0,24
	ISR „Essstörungsskala“	0,08	0,05	0,00
Woche 1	OQ-45 „Gesamtskala“	-	0,07	0,16
	OQ-45 Skala „Symptombelastung“	-	0,08	0,17
	OQ-45 Skala „Zwischenmenschliche Beziehungen“	-	0,02	0,09
	OQ-45 Skala „Soziale Integration“	-	0,05	0,14
	ISR „Gesamtskala“	-	0,01	0,08
	ISR „Depressionsskala“	-	0,03	0,14
	ISR „Angstskala“	-	-0,05	0,01
	ISR „Zwangsskala“	-	0,02	0,01
	ISR „Somatisierungsskala“	-	0,07	0,17
	ISR „Essstörungsskala“	-	-0,03	-0,08
Woche 2	OQ-45 „Gesamtskala“	-	-	0,09
	OQ-45 Skala „Symptombelastung“	-	-	0,09
	OQ-45 Skala „Zwischenmenschliche Beziehungen“	-	-	0,06
	OQ-45 Skala „Soziale Integration“	-	-	0,08
	ISR „Gesamtskala“	-	-	0,07
	ISR „Depressionsskala“	-	-	0,11
	ISR „Angstskala“	-	-	0,06
	ISR „Zwangsskala“	-	-	-0,01
	ISR „Somatisierungsskala“	-	-	0,09
	ISR „Essstörungsskala“	-	-	-0,05

balen Kriterien signifikant gebessert, wobei bereits von Woche 0 bis Woche 1 die Verbesserungen signifikant wurden. Über die Verbesserung von Woche 0 zu Woche 1 hinaus besserten sich von Woche 1 zu Woche 3 noch die globale Symptombelastung (OQ-45-Skala „Symptombelastung“) und die psychische Gesamtbelastung (OQ-45 „Gesamtskala“) signifikant. Obwohl in manchen Skalen nur die Verbesserung von Woche 0 zu Woche 1 statistische Signifikanz erreichte, stiegen die Effektstärken aller OQ-45-Skalen und der ISR „Gesamtskala“ doch pro Behandlungswoche an, so dass jede der hier untersuchten Behandlungswochen zur globalen Effektivität der psychosomatischen Behandlung beitrug. Die Effektstärken, die in der OQ-45 „Gesamtskala“ für die ersten Behandlungswochen gefunden wurden, waren vergleichbar mit den Effektstärken, wie sie aus den von Puschner et al. (2009) berichteten Ergebnissen zur stationären psychiatrischen Behandlung errechnet werden können (Psychiatrie: Woche 0 zu Woche 1 $d=0,22-0,30$; Woche 1 zu Woche 2: $d=0,02-0,09$; unsere Studie: Woche 0 zu Woche 1 $d=0,25$; Woche 1 zu Woche 2: $d=0,07$). Auch war die Effektstärke für die OQ-45 „Gesamtskala“ von Woche 0 zu Woche 2 von $d = 0,32$ vergleichbar mit der Effektstärke von $d = 0,33$, die sich aus den statistischen Angaben bei Percevic et al. (2006) von Aufnahme zur Zwischenmessung (18 Tage nach Aufnahme) errechnen lässt. Deutlich höher waren die Effektstärken in den OQ-45 Skalen in unserer Studie als die, die sich bei Bauer (2004) zeigten: Von Aufnahme zur Zwischenmessung (3–5 Wochen nach Aufnahme) lagen bei Bauer (2004) die Effektstärken zwischen $d=0,0$ (OQ-45 Skala „Symptombelastung“) und $d=0,2$ (OQ-45 Skala „Soziale Integration“).

Anders als das Phasenmodell therapeutischer Veränderungen von Howard et al. (1993) vermuten lassen würde, verbesserten sich in dieser Untersuchung psychische Symptome nicht vor interpersonellen und sozialen Problemen. Dies kann einerseits daran liegen, dass unsere statistische Analyse der Veränderungsraten nicht wie bei Howard et al. (1993) auf individuellen Bes-

serungsraten basierte: Höchstwahrscheinlich zeigen sich unterschiedliche Behandlungsverläufe, je nachdem, ob auf statistische Signifikanz geprüft wird, Effektstärken oder andere statistische Methoden verwendet werden, z.B. der Reliable Change Index (RCI) bzw. das Konzept der klinisch signifikanten Verbesserungen (Jacobson & Truax, 1991) oder die Methode der prozentualen Symptomreduktion (Hiller et al. 2011): Melchior (2011) z.B. fand, dass sich nach der ersten Behandlungswoche „nur“ 14% klinisch signifikant verbesserten, während es zur siebten Behandlungswoche mit 53% deutlich mehr Patienten waren, die dieses Kriterium erreichten. Andererseits wäre auch denkbar, dass für die multimodale und multiprofessionelle stationäre psychosomatische Behandlung ein anderes Phasenmodell geeigneter sein könnte als das von Howard et al. (1993), das auf einem Setting typisch für eine psychotherapeutische Privatpraxis basiert.

Die in dieser Arbeit analysierten spezifischen Kriterien (ISR-Syndromskalen) veränderten sich zwischen Woche 0 und Woche 3 differenziell: Die Symptombelastung durch depressive Symptome (ISR „Depressionsskala“) besserte sich wie auch die Symptombelastung durch Angstsymptome (ISR „Angstskala“) bereits signifikant zwischen Woche 0 und Woche 1. Die Symptombelastung durch Somatisierungssymptome (ISR „Somatisierungsskala“) dagegen zeigten erstmals signifikante Verbesserungen von Woche 0 bzw. Woche 1 bis Woche 3. Dass sich Somatisierungssymptome später als depressive Symptome bzw. Angstsymptome veränderten, könnte teilweise daran liegen, dass besonders bei Krankheitsbildern mit Somatisierungssymptomen in der ersten Therapiephase die Erarbeitung eines adäquaten Krankheitsmodells im Vordergrund steht (Hausteiner-Wiehle et al. 2013). Die Symptombelastung durch Zwangssymptome bzw. Essstörungssymptome erzielte hingegen noch keine signifikante Verbesserung zwischen Woche 0 und Woche 3. Hierbei könnten zwei Aspekte zur Erklärung herangezogen werden: Erstens haben, wie bereits erwähnt, nur wenige Patienten der Stichprobe Zwangs- bzw. Essstörungen diagnostiziert bekommen, so dass

während der Aufnahmewoche auf der ISR „Zwangs-“ bzw. ISR „Essstörungsskala“ die geringsten spezifischen Symptombelastungen vorhanden waren. Dies hat zur Folge, dass sich auf diesen beiden Skalen aufgrund geringerer Ausgangswerte auch nur geringere Verbesserungen ergeben können. Zweitens haben von Heymann et al. (2003) gezeigt, dass Patienten mit Zwangsstörungen und Patienten mit Essstörungen eine längere Behandlungsdauer in der stationären Psychosomatik haben als Patienten mit depressiver Störung, Angststörung oder somatoformer Störung. Es

könnte vermutet werden, dass Veränderungen in den ersten drei Wochen nach Aufnahme bei Essstörungen- und Zwangssymptomen weniger erwartet werden können als bei depressiven Symptomen bzw. Angst- oder Somatisierungssymptomen. Weitere Untersuchungen sind von Nöten, um festzustellen, ab welcher Behandlungswoche sich Symptome einer Zwangsstörung bzw. Essstörung verbessern. In dieser Arbeit waren aufgrund der kurzen Behandlungsdauern in den beteiligten Kliniken nicht genügend Daten verfügbar, um Veränderungen jenseits der dritten Woche nach Aufnahme untersuchen zu können.

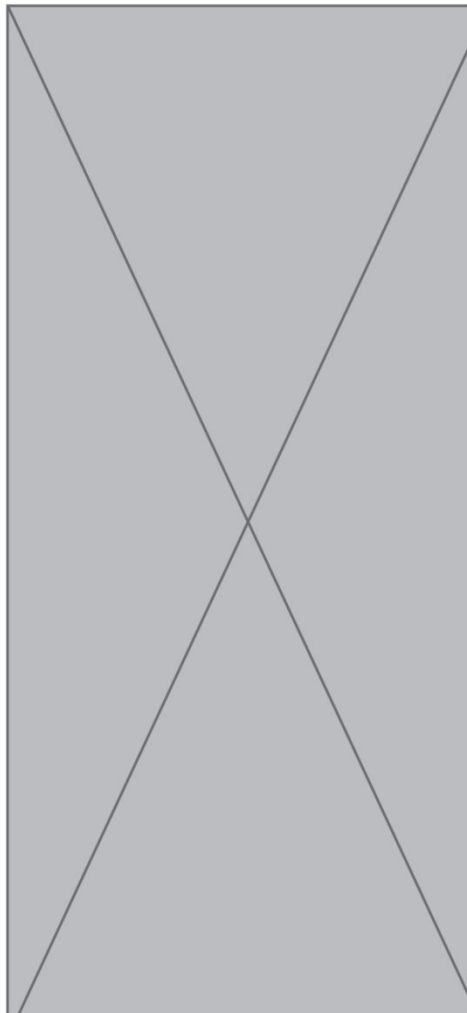
Eine Einschränkung der vorliegenden Studie sind die Abstriche bei der internen Validität bedingt durch die fehlende Kontrollgruppe. Damit können die festgestellten Effekte nicht eindeutig auf die psychosomatische Therapie bezogen werden. Demgegenüber ist die externe Validität dieser Studie durch die Realisierung unter den Bedingungen der Routinepraxis relativ hoch. Dabei sollte noch bedacht werden, dass die externe Validität der Ergebnisse zur Änderungsdiagnostik möglicherweise durch Stichprobenselektionseffekte beeinflusst wurde: Die Drop-Out-Analyse hat gezeigt, dass sich die Patienten der Stichprobe zur Änderungsdiagnostik in einigen Variablen von den Drop-Outs unterscheiden. Entsprechend wäre eine Replikation dieser Studie mit größerer Stichprobe insbesondere mit mehr Patienten mit Zwangs- bzw. Essstörungen sinnvoll.

Zusammengefasst zeigte diese Untersuchung, dass sich während der Aufnahmewoche die Mehrheit der Patienten in psychodynamisch-orientierter stationärer psychosomatischer Therapie als belastet einschätzte und sich unterschiedliche Beschwerden von der Aufnahmewoche bis zu den folgenden drei Behandlungswochen differenziell veränderten.

Literatur online unter www.pdp-online.info

Korrespondenzadresse

Dipl. Psych. Thomas Probst
Klinische Psychologie und Psychotherapie
Universität Regensburg
Universitätsstr. 31
93053 Regensburg
E-Mail: Thomas.Probst@psychologie.uni-regensburg.de



Literatur

- Bauer S (2004). Ergebnismonitoring und Feedback: Mittel zur Optimierung stationärer Psychotherapie? Dissertation zur Erlangung des akademischen Grades Doktor der Sozialwissenschaften in der Fakultät für Sozial- und Verhaltenswissenschaften der Eberhard-Karls-Universität Tübingen.
- Brandt WA (2010). Kreuzvalidierung des ICD-10-Symptom-Rating-(ISR)-Selbstbeurteilungsfragebogens an 969 depressiven Patienten mit der SCL-90-R und dem BDI I. Dissertation, Fakultät für Medizin, Universität Regensburg.
- Fischer F, Tritt K, Klapp BF, Fliege H (2010). Faktorenstruktur und psychometrische Eigenschaften des ICD-10-Symptom-Rating (ISR) an Stichproben psychosomatischer Patienten. *Psychotherapie Psychosomatik Medizinische Psychologie* 60, 307–315.
- Fischer F, Schirmer N, Tritt K, Klapp BF, Fliege H (2011). Retest-Reliabilität und Änderungssensitivität des ICD-10-Symptom-Rating (ISR) in verschiedenen Stichproben. *Psychotherapie Psychosomatik Medizinische Psychologie* 61, 162–169.
- Haase M, Frommer J, Franke G-H, Hoffmann T, Schulze-Muetzel J, Jäger S et al (2008). From symptom relief to interpersonal change: treatment outcome and effectiveness in inpatient psychotherapy. *Psychotherapy Research* 18, 615–624.
- Hausteiner-Wiehle C, Schäfer R, Sattel H, Ronel J, Hermann M, Häuser W et al (2013). Neue Leitlinie zu funktionellen und somatoformen Störungen. *Psychotherapie Psychosomatik Medizinische Psychologie* 63, 26–31.
- Hiller W, Schindler A, Lambert MJ (2011). Defining response and remission in psychotherapy research: a comparison of the RCI and the method of percent improvement. *Psychotherapy Research* 22, 1–11.
- Howard KI, Lueger RJ, Maling MS, Martinovich Z (1993). A phase model of psychotherapy outcome: causal mediation of change. *Journal of Consulting and Clinical Psychology* 61, 678–685.
- Howard KI, Moras K, Brill PL, Martinovich Z, Lutz W (1996). Evaluation of psychotherapy: efficacy, effectiveness, and patient progress. *American Psychologist* 51, 1059–1064.
- Jacobson NS, Truax P (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 59, 12–19.
- Kopta SM, Howard KI, Lowry JL, Beutler LE (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology* 62, 1009–1016.
- Lambert MJ (2012). Helping clinicians to use and learn from research-based systems: the OQ-analyst. *Psychotherapy* 49, 109–114.
- Lambert MJ, Kahler M, Harmon C, Shimokawa K, Burlingame G (2011). Administration and scoring manual for the Outcome Questionnaire-45.2. Orem, UT: American Professional Credentialing Services.
- Lambert MJ, Hannöwer W, Nisslmüller K, Richard M, Kordy H (2002). Fragebogen zum Ergebnis von Psychotherapie: zur Reliabilität und Validität der deutschen Übersetzung des Outcome Questionnaire 45.2 (OQ-45.2). *Zeitschrift für Klinische Psychologie und Psychotherapie* 31, 40–47.
- Liebherz S, Rabung S (2013). Wirksamkeit psychotherapeutischer Krankenhausbehandlung im deutschsprachigen Raum: Eine Meta-Analyse. *Psychotherapie Psychosomatik Medizinische Psychologie* 63, 355–364.
- Melchior H (2011). Vorhersage des längerfristigen Behandlungserfolgs anhand von Symptomverläufen und der Dauer stationärer Psychotherapie. Dissertation zur Erlangung der Würde des Doktors der Philosophie am Fachbereich Psychologie, Fakultät für Erziehungswissenschaft, Psychologie und Bewegungswissenschaft der Universität Hamburg.
- Percevic R, Lambert MJ, Kordy H (2006). What is the predictive value of responses to psychotherapy for its future course? Empirical explorations and consequences for outcome monitoring. *Psychotherapy Research* 16, 364–373.
- Probst T, Lambert MJ, Loew TH, Dahlbender RW, Göllner R, Tritt K (2013). Feedback on patient progress and clinical support tools for therapists: improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice. *Journal of Psychosomatic Research* 75, 255–261.
- Probst T, Lambert MJ, Dahlbender RW, Loew TH, Tritt K (in press). Providing patient progress feedback and clinical support tools to therapists: Is the therapeutic process of patients on-track to recovery enhanced in psychosomatic in-patient therapy under the conditions of routine practice? *Journal of Psychosomatic Research*.
- Puschner B, Schöfer D, Knaup C, Becker T (2009). Outcome management in in-patient psychiatric care. *Acta Psychiatrica Scandinavica* 120, 308–319.
- Seidel M (2013). Vergleich der psychometrischen Fragebögen ICD-10-Symptom-Rating (ISR) und der Langform des Patient-Health-Questionnaire (PHQ). Dissertation, Fakultät für Medizin, Universität Regensburg.
- Shimokawa K, Lambert MJ, Smart DW (2010). Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology* 78, 298–311.
- Steffanowski A, Löschmann C, Schmidt J, Wittmann WW, Nübling R (2007). Metaanalyse der Effekte stationärer psychosomatischer Rehabilitation. Bern: Huber.
- Tritt K, von Heymann F, Zaudig M, Loew TH, Söllner W, Fischer HF, Bühner M (2010a). Das ICD-10-Symptom-Rating (ISR): Zusammenhänge mit der SCL-90-R. *Klinische Diagnostik und Evaluation* 1, 22–37.
- Tritt K, von Heymann F, Zaudig M, Söllner W, Klapp B, Loew TH, Bühner M (2010b). Der Fragebogen ICD-10-Symptom-Rating (ISR). Kurzdarstellung der Normierung. Institut für Qualitätsentwicklung in der Psychotherapie und Psychosomatik.
- Tritt K, von Heymann F, Zaudig M, Zacharias I, Söllner W, Loew TH (2008). Entwicklung des Fragebogens „ICD-10-Symptom-Rating“ (ISR). *Zeitschrift für Psychosomatische Medizin und Psychotherapie* 54, 420–429.
- von Heymann F, Zaudig M, Tritt K et al (2003). Die diagnosebezogene Behandlungsdauer in der Psychosomatischen und Psychotherapeutischen Medizin: eine homogene Größe? Erste Ergebnisse der Multicenter-Basisdokumentation (Psy-BaDo-PTM) als Grundlage qualitätssichernder Maßnahmen in der stationären Psychosomatik. *Praxis Klinische Verhaltensmedizin und Rehabilitation* 62, 209–221.
- Wittchen H-U, Jacobi F (2012). Was sind die häufigsten psychischen Störungen in Deutschland? URL: http://www.rki.de/DE/Content/Gesundheitsmonitoring/Studien/Degs/degs_w1/Symposium/degs_psy-chische_stoerungen.pdf?__blob=publicationFile Letzter Zugriff: 22.11.2013

III. 4. Publikation 4

Probst, T., Lambert, M. J., Loew, T. H., Dahlbender, R. W., & Tritt, K. (under re-review). Extreme deviations from expected recovery curves and their associations with therapeutic alliance, social support, motivation, and life events in psychosomatic in-patient therapy. *Psychotherapy Research*

Title: Extreme deviations from expected recovery curves and their associations with therapeutic alliance, social support, motivation, and life events in psychosomatic in-patient therapy

Running head: Extreme deviations from ERCs and the ASC scales

Abstract:

Objectives: Within the clinical support tools (CST) of the OQ-Analyst, the "Assessment for Signal Cases" (ASC) evaluates the therapeutic alliance, social support, motivation, and life events. We investigated whether the ASC covers domains of importance in treatment weeks with extreme deviations from expected recovery curves (ERCs).

Methods: Psychosomatic in-patients were monitored weekly with the ASC and the "Outcome Questionnaire" (OQ-45). The ERCs of the OQ-45 empirical algorithm were used to define treatment weeks with extreme positive deviations (EPD), extreme negative deviations (END), or without extreme deviations (NO). Associations between the ASC scales and EPD as well as END were analyzed by multilevel models.

Results: While each ASC scale was positively associated with EPD, only the social support and life events scales were negatively related to END.

Conclusions: CSTs prioritizing social support and life events might be more effective in preventing treatment failure.

Keywords: Monitoring; considerable change; psychotherapeutic process; psychosomatics

Introduction:

In patient-focused research (Howard, Moras, Brill, Martinovich, & Lutz, 1996), the individual patient's progress is continuously monitored with patient reported outcome measures (Greenhalgh, 2009) throughout treatment. By tracking patient progress during psychotherapies, time intervals with considerable positive or negative change can be identified for individual patients.

Different methods have been applied in psychotherapy research to operationalize considerable change with no agreed-upon standard procedure. For example, the "Reliable Change Index" (RCI; Jacobson, & Truax, 1991) can be applied to define reliable positive and negative change. The method of percent improvement (Hiller, Schindler, & Lambert, 2013) was suggested to evaluate response and remission. A definition proposed by Tang and DeRubeis (1999) should operationalize sudden dramatic change in the positive (sudden gains: e. g. Tang, & DeRubeis, 1999) or negative (sudden losses: e. g. Lutz et al., 2013) direction. In another method to identify considerable positive or negative change the progress of a given patient is compared with his or her expected recovery curve (ERC) derived from a norm group consisting of patients with characteristics comparable to the patient in question (e. g. comparable levels of distress at intake; Finch, Lambert, & Schaalje, 2001; Newnham, Hooke, & Page, 2010). Extreme positive or negative deviations of the individual patient's progress from his or her ERC can be considered as either considerable positive or considerable negative change depending on the direction of change.

A variety of mechanisms associated with considerable change have been critically discussed especially for sudden gains in cognitive-behavior therapy (CBT) of depression: Some authors advocate the possible causal role of specific treatment mechanisms (Andrusyna, Luborsky,

Pham, & Tang, 2006; Tang, & DeRubeis, 1999; Tang, DeRubeis, Beberman, & Pham, 2005).

Others have speculated that common factors - i.e. factors viewed as active in all types of psychotherapy - or client variables are crucial, particularly for considerable positive change in the early phases of treatment (Ilardi, & Craighead, 1994; Lambert, 2005).

A central question to be addressed is which factors should be considered as relevant with regard to considerable change. Taking quantitative findings into account, Asay and Lambert (1999) estimated that the following four general factors are each more or less relevant for psychotherapy outcome: The contribution to treatment response was considered to amount to 40% for client and extratherapeutic factors, 30% for therapeutic relationship factors, 15% for hope and expectancy factors, and 15% for models and technique factors. A recent meta-analysis provided empirical evidence for these estimations regarding the psychotherapy outcome of depressive disorders (Cuijpers et al., 2012). Building on these findings, the self-report questionnaire “Assessment for Signal Cases“ (ASC; Lambert et al., 2007) was developed: Within the ASC, the extratherapeutic factors are operationalized with two scales (social support scale and life events scale), the therapeutic relationship factors are evaluated with a therapeutic alliance scale, and the hope and expectancy factors are covered by a motivation scale. The models and technique factors are not included in the ASC, but other questionnaires are available to assess these factors, e. g. the “Multitheoretical List of Therapeutic Interventions“ (MULTI; McCarthy, & Barber, 2009).

Providing therapists with feedback on the ASC results constitutes the first part of the clinical support tools (CSTs) of the monitoring and feedback software OQ-Analyst (Lambert, 2012): When a patient deviates negatively from his or her ERC by at least one standard deviation (such a patient is considered at-risk of treatment failure; Lambert, Whipple et al., 2002; Spielmans, Lambert, & Smart, 2006), the therapist should administer the ASC to this patient and examine the ASC scores (pertaining to the last week) to identify possible problem areas

that could account for the at-risk classification. The second part is recommended when patients show problematic scores on ASC scales: This part consists of a decision tree supporting therapists in organizing problem-solving strategies as well as in selecting interventions to improve problematic ASC domains (Lambert, et al., 2007).

When therapists receive CSTs in addition to patient progress feedback, the outcome of patients at risk of treatment failure has been shown to improve in counseling (Harmon et al., 2007; Slade, Lambert, Harmon, Smart, & Bailey, 2008; Whipple et al., 2003), outpatient psychotherapy (Simon, Lambert, Harris, Busath, & Vazquez, 2012), outpatient substance abuse treatment (Crits-Christoph et al., 2012), and psychosomatic in-patient therapy (Probst et al., 2013).

Although the ASC comprises domains shown to be relevant for the prevention of treatment failure as well as for psychotherapy outcome / progress (therapeutic alliance: e. g. Horvath, Del Re, Flückiger, & Symonds, 2011; social support: e. g. Bankoff, & Howard, 1992; motivation: e. g. Pelletier, Tuson, & Haddad, 1997; and life events: e. g. Sexton, 1996), the influence of those domains on considerable change during psychotherapy remains unclear. Previous studies examining associations between those domains and considerable change either on a patient-level or an assessment-level showed inconsistent results:

The patient-level studies compared patients who experienced considerable change at least at one assessment point during the entire course of psychotherapy against patients without any considerable change: It has been shown that neither the therapeutic alliance of the first psychotherapy session (Hardy et al., 2005), nor social support at the time of pretreatment (Hopko, Robertson, & Carvalho, 2009; Keller, Feeny, & Zoellner, 2014), nor life events during the course of psychotherapy (Hardy et al., 2005) differed between patients with and without at least one sudden gain. Furthermore, the therapeutic alliance during the entire

course of psychotherapy did not differ between patients with at least one sudden gain assessment and patients with at least one sudden loss assessment (Lutz et al., 2013).

As the presence or absence of considerable change can be investigated at each available assessment point (this does not hold for the intake assessment, since at least two assessments are necessary to obtain a change score), a patient reaching considerable change at least at one assessment can also provide assessments not coded as considerable change during the course of treatment. The instability of considerable change has for instance been shown for sudden gains (Lutz, Bachmann, Tschitsaz, Smart, & Lambert, 2007). Therefore, the assessment-level studies investigated how variables of interest (here the ASC domains) are related to individual assessments coded as considerable change. Those studies revealed that patients with depressive disorders rate the therapeutic alliance higher not before but after sudden gain assessments (Tang, & DeRubeis, 1999) and that higher motivation preceded assessments coded as sudden gains in patients with eating disorders (Cavallini, & Spangler, 2013). Moreover, independent observers considered therapists as more distanced and impatient in sudden loss sessions but more empathetic in sessions without considerable change in outpatient psychotherapies (Tschitsaz-Stucki, & Lutz, 2009). Furthermore, the patients' motivation as well as their amount of life events were rated higher in sessions without considerable change, but their expectations were considered to be more realistic in sessions coded as sudden gains (Tschitsaz-Stucki, & Lutz, 2009).

The cited studies did not investigate associations between considerable change and the ASC domains in psychosomatic in-patient therapy. Therefore, the present study examined with psychosomatic in-patients whether the ASC scales are of importance during treatment weeks with considerable change. We hypothesized that the weekly ASC scale scores are 1) positively related to treatment weeks with considerable positive change and 2) negatively associated with treatment weeks coded as considerable negative change.

Methods:

Questionnaires:

In this study, two questionnaires were administered to the patients simultaneously on a weekly basis:

Assessment for Signal Cases (ASC): The self-report questionnaire ASC (Lambert et al., 2007) contains 40 items which are rated on a five point Likert scale. The ratings pertain to the previous week. The ASC covers four domains: Therapeutic alliance (11 items, range: 11-55), social support (11 items, range: 11-55), motivation (9 items, range: 9-45), and life events (9 items, range: 9-45). The higher the ASC scores, the better the therapeutic alliance, social support, motivation, and life events. For this study, the ASC was translated into German using the back-translation method as well as cognitive debriefings with 10 in-patients (Wild et al., 2005). The internal consistency coefficients (Cronbach's Alpha) of the scales of the German ASC were calculated with the ASC assessments at intake resulting in the following Cronbach's Alphas: Therapeutic alliance scale: 0.89; social support scale: 0.76; motivation scale: 0.78; life events scale: 0.71.

Outcome Questionnaire (OQ-45): The OQ-45 (Lambert, Hannöver, Nisslmüller, Richard, & Kordy, 2002; Lambert, Kahler, Harmon, Burlingame, & Shimokawa, 2011) is a self-report questionnaire which comprises 45 items rated on a five point Likert scale. All ratings refer to the last week. In this paper, only the total scale of the OQ-45 (min: 0, max: 180) was used to measure the general distress level due to psychological problems and to define considerable change (see below). For the total scale of the German version of the OQ-45, the internal consistency reached 0.93 and the retest reliability amounted to $r = 0.88$ (Lambert, Hannöver, et al., 2002). The higher the score in the OQ-45, the worse the impairment.

Study design:

Patients of two psychosomatic in-patient clinics were monitored over the course of their treatment. They were asked to fill out the German paper and pencil versions of the OQ-45 and the ASC once a week (between Friday and Sunday) during their stay. While all patients of the Psychosomatics Department of the Hospital in Donaustauf, Germany - treated from 10/01/2010 to 07/06/2012 - were asked to participate in the study, only patients with private health care of the Psychosomatic Hospital “Am Schönen Moos” in Bad Saulgau, Germany - treated from 01/16/2012 to 04/06/2012 - were encouraged to take part in the study. In both clinics, a multiprofessional team (consisting mainly of psychologists, physicians, and nurses) provides a multimodal treatment with more than 24 treatment hours a week (including at least two individual psychotherapy sessions per week, group psychotherapy, relaxation and mindfulness training as well as art, music, and dance therapy and - if required - crisis intervention, visitations by nurses, medical consultations).

In our overall project, patients were randomly assigned either to the control group or to the experimental group. For patients of the experimental group, the therapists responsible for individual therapy received continuous patient progress (OQ-45) feedback as well as feedback on the ASC results. This paper analyses the scores of both control group and experimental group patients. The feedback effects were reported in other publications (Probst et al., 2013; Probst, Lambert, Dahlbender, Loew, & Tritt, 2014). The study was approved by the Ethics Committee of the University Clinic of Regensburg.

Definition of considerable change:

Using the monitoring and feedback software OQ-Analyst (Lambert, 2012), the “empirical algorithm” of the OQ-45 was applied to define the presence or absence of considerable change on a weekly basis.

The “empirical algorithm” provides an ERC for each patient as a function of the patient’s intake OQ-45 score. The ERC of a given patient is based on a norm group of at least 220 patients with OQ-45 intake scores comparable to the patient in question (Finch, Lambert, & Schaalje, 2001). Since the intake score and the change from session 1 to 3 has been shown to account for 40% of the outcome variance and adding other variables (e. g. diagnosis, therapy schools) did not prove itself useful in explaining more variance in the final outcome (Lambert, 2013), only the intake score functions as the baseline for the ERC.

Each treatment week following the intake assessment, the OQ-45 score of a given patient was compared with his or her ERC. When the OQ-45 score of a patient deviated positively from the patient’s ERC and this score ranged within the top 10% of the norm group used to generate the ERC (“blue signal patients”; Lambert, 2012), the week was coded as treatment week with an extreme positive deviation (EPD). When an OQ-45 score deviated negatively by at least one standard deviation from the patient-specific ERCs (bottom 16% of the norm group), the week was coded as treatment week with an extreme negative deviation (END). This criterion has also been used to define patients at risk of treatment failure (“not-on-track” patients) in previous studies (Lambert, Whipple et al., 2002; Spielmans et al., 2006). Weeks with OQ-45 scores not deviating as described from the patient-specific ERCs were classified as treatment weeks without extreme deviations (NO).

Statistical analysis:

The statistical analysis was done with SPSS 21. Means (M), standard deviations (SD), and percentages (%) were calculated as descriptive statistics.

For investigating associations between the ASC scales and considerable change, the following analyses were done with the weekly assessments (it should be kept in mind that the assessments taken during the intake week were only used to determine the subject's ERC and are hence not included in any further calculations).

For examining treatment weeks with EPD, the variable EPD was created for each treatment week: If an OQ-45 assessment was classified as EPD, this new variable was coded as 1 and if an OQ-45 assessment was classified as END or NO, this variable was coded as 0.

To analyze treatment weeks with END, the variable END was also created for each treatment week: If an OQ-45 assessment was classified as END, this variable was coded as 1 and if an OQ-45 assessment was classified as EPD or NO, this variable was coded as 0.

Two multilevel models for dichotomous outcomes (Heck, Thomas, & Tabata, 2012) were performed to study associations between the ASC scales and EPD or END. Both models were binary logistic models within the "Generalized Estimating Equations" (GEE) in SPSS.

Multilevel models were chosen due to the nested structure of the data: Assessment points as level 1 nested within patients as level 2. In GEE, patients were therefore designated as subject variable and assessment points as within-subject variable. The therapists and clinics were not added as levels to the multilevel models due to the confidentiality agreement with the therapists and clinics.

Model 1 investigated associations between the ASC scales and EPD. EPD was the binary outcome: 1 (treatment week coded as EPD) was set as the answer and 0 (treatment week coded as either END or NO) as the reference.

Model 2 analyzed associations between the ASC scales and END. In this model, END was the binary outcome: Again, 1 (treatment week coded as END) was defined as the answer and 0 (treatment week coded as either EPD or NO) as the reference.

In both models, the four ASC scales (therapeutic alliance, social support, motivation, life events) were added as covariates (time-dependent: the ASC scale scores always pertained to the same treatment week as the EPD, END, or NO classifications). The four covariates were grand-mean centered, since the ASC scales do not have a meaningful 0 value. In GEE, the main-effects setting was selected to obtain a main-effects term for each of the four covariates. Moreover, the following settings were chosen for both models: Covariance matrix=robust estimator, analysis type=type 3, confidence interval=95%, chi-square statistics=Wald, log quasi likelihood function=full.

Regarding the working correlation matrix structure, we tried three structures: Unstructured, exchangeable (compound symmetry structure), and autoregressive. Of those three structures, the one was chosen as the final structure of model 1 / model 2 with which the model produced convergence and the smallest QIC (quasi likelihood under independence model).

All tests were performed two-tailed with a p-value of ≤ 0.05 . No corrections for multiple testing (e. g. Bonferroni) were applied due to the explorative nature of our study (Abt, 1987).

Participants:

The patients included in this paper had to fulfill two criteria:

First, they had to provide an intake OQ-45 score and at least one more OQ-45 score. We excluded all other patients, since the intake OQ-45 scores are necessary to plot the ERCs of the empirical OQ-45 algorithm and only assessments following the intake assessments can be

classified as either EPD, END, or NO. Second, patients were only included if they completed the OQ-45 and the ASC simultaneously at least at one assessment point following intake, since at least one simultaneous OQ-45 and ASC assessment is necessary to study associations between considerable change and the ASC scales.

Of all 436 registered patients, 271 patients (62%) fulfilled these inclusion criteria. The excluded 165 patients either did not provide an OQ-45 score for the intake week, since they were too distressed or not motivated to take part in the study (65% of 165) or they did not complete the OQ-45 and ASC simultaneously at least at one assessment point after the intake assessment (35% of 165).

The characteristics of the 271 patients included in this study are presented in table 1. The diagnoses were made by the clinic teams according to chapter V of the ICD-10 (WHO, 1992). Taking psychological comorbidity into account, all diagnoses of the included patients (not only the main diagnosis) are reported in table 1.

--- Table 1 here ---

Results:

Treatment weeks with EPD, END, or NO:

Altogether the included patients provided 791 weekly OQ-45 assessments following the intake assessment that were classified as either EPD, END, or NO: Of all these assessments, 178 (22.5%) were classified as EPD, 75 (9.5%) as END, and the rest (538, 68.0%) as NO. The 178 weeks with EPD belonged to 94 patients (34.7% of 271), the 75 weeks with END to 38 patients² (14.0% of 271), and the 538 weeks with NO to 231 patients (85.2% of 271).

Importance of the ASC scales during treatment weeks with EPD:

The multilevel model used to study associations between the ASC scales and EPD did not converge with the unstructured working correlation matrix structure. However, the model converged with the autoregressive structure and the exchangeable structure. The autoregressive structure showed a slightly smaller QIC (688.92) than the exchangeable structure (689.60). Therefore, the results of the model with the autoregressive structure are reported in table 2.

--- Table 2 here ---

² The number of patients with END amounted to 38 in this paper and differs from the 43 patients with END analyzed by Probst et al. (2013) since 5 of those 43 patients did not complete the ASC in their treatment weeks classified as END.

As can be seen in table 2, each ASC scale was significantly positively ($\beta > 0$, $\text{Exp}(\beta) > 1$, $p < 0.05$) associated with EPD when the other ASC scales were hold constant at their grand-means.

Relevance of the ASC scales during treatment weeks with END:

The multilevel model performed to analyze associations between the ASC scales and END converged with the unstructured, the autoregressive, as well as with the exchangeable working correlation matrix structure. The autoregressive structure produced a smaller QIC (436.70) than the exchangeable structure (443.92) and the unstructured structure (444.61). Table 3 thus presents the results of the model with the autoregressive structure.

--- Table 3 here ---

Table 3 illustrates that only the ASC social support scale and the ASC life events scale were significantly negatively ($\beta < 0$, $\text{Exp}(\beta) < 1$, $p < 0.05$) related to END when the other ASC scales were hold constant at their grand-means.

Discussion:

This study investigated in two psychosomatic in-patient clinics whether the domains of the questionnaire “Assessment for Signal Cases” (ASC) - therapeutic alliance, social support, motivation, and life events - are of importance when patients deviate extremely from their expected recovery curves (ERCs). The ERCs of the OQ-45 empirical algorithm (Finch et al., 2001) were applied to define treatment weeks with extreme positive deviations (EPD), extreme negative deviations (END) or without extreme deviations (NO). The ASC and the OQ-45 were administered to patients simultaneously on a weekly basis during their in-patient stay so that the ASC scales pertained to the weeks coded as either EPD, END, or NO.

The percentage of patients found with at least one END (14.0%) occurring during the entire course of treatment was within the range found in previous studies that also used the OQ-45 algorithms to operationalize considerable negative change (Amble et al., 2014; Crits-Christoph et al., 2012; Shimokawa, Lambert, & Smart, 2010; Simon et al., 2012: 11% - 56%). Although the method used to operationalize considerable positive change in this study (OQ-45 algorithm) differs from previous research (which mostly used the sudden gain definition proposed by Tang, & DeRubeis, 1999), our result of 34.7% patients with at least one EPD also lies within the 11% - 50% interval of sudden gain patients found in other investigations (Lutz et al., 2013). Of all assessments taken during the weeks following intake, about twenty-three percent were classified as EPD and ten percent were coded as END. To investigate the relevance of the ASC scales during treatment weeks with considerable change (EPD, END), multilevel models for dichotomous outcomes were applied. Across all treatment weeks, all four ASC scales were significantly positively related to EPD. However, significant negative associations with END were found only for the ASC social support scale and the ASC life

events scale, but neither for the ASC motivation scale nor for the ASC therapeutic alliance scale.

Taking the sheltered in-patient setting of our sample into account, the importance of extratherapeutic factors for EPD and END was somewhat surprising for us. The usefulness of the extratherapeutic scales of the ASC - especially of the social support scale - has been questioned for the psychosomatic in-patient setting (Probst et al., 2013). However, this study highlights the relevance of extratherapeutic factors (especially for END) even in psychosomatic in-patient therapy. Social support could for instance be important for considerable change in in-patient settings, since in-patients might experience more or less social support from other in-patients. Social support could therefore be at least partially related to cohesiveness, a curative factor proven to influence patient progress in psychosomatic in-patient therapy (Dinger, & Schauenburg, 2010). Due to the possible overlap of social support and cohesiveness in in-patient settings, social support might be both an extratherapeutic factor and a factor within therapy in in-patient settings. Life events could be associated with considerable change as some in-patients undergo stress tests (tests of momentary coping abilities) on the weekends during their in-patient stay. Moreover, as in-patients live some weeks of their lives in the clinics, they may experience life events (e. g. receiving bad news from relatives or work) also during their in-patient stay.

Our significant associations between considerable positive change and social support as well as life events seem to contradict previous studies which applied the sudden gain definition to operationalize considerable positive change: The relationship between social support and sudden gains did not attain statistical significance in the studies of Hopko et al. (2009) and Keller et al. (2014). However, these studies assessed social support only at pretreatment and did not analyze the association between individual assessments and social support over the course of treatment. Furthermore, although Hardy et al. (2005) assessed life events on a

session-by-session basis, they failed to show a significant relationship between sudden gains and life events. Besides the different methods used to operationalize considerable positive change, two more differences between our investigation and the study of Hardy et al. (2005) should be considered and could also explain the ambiguous results: First, Hardy et al. (2005) compared life events over the course of treatment between patients with at least one sudden gain against patients without any sudden gain. In contrast, we analyzed life events pertaining to weeks with and without considerable positive change. Second, Hardy et al. (2005) dichotomized the life events score (good vs. bad life event scores), we in turn used a life events scale score.

The impact of motivation as well as of the therapeutic alliance could have turned out being stronger for EPD than for END in our study because we applied a more rigorous criterion for defining EPD than for defining END. The top 10% of the norm group were the reference for our EPD classification, and the bottom 16% of the norm group for our END classification. The different criteria applied to operationalize EPD and END could be viewed as a limitation of this study. Wanting to enhance the comparability to other studies, we decided to use the bottom 16% to define END since patients scoring in this range were considered “not-on-track” patients, signal cases, or patients at risk of treatment failure in previous research (Lambert, Whipple et al., 2002; Spielmans et al., 2006). The top 10% were used define EPD since the OQ-Analyst labels patients scoring in this range as “blue signal patients” (Lambert et al., 2012).

Another limitation refers to our examination of only four domains (therapeutic alliance, social support, motivation, life events) in weeks with and without considerable change. Other domains shown to influence patient progress in psychosomatic in-patient therapy (e. g. Dinger, & Schauenburg, 2010; Mander et al., 2013; 2014) were not assessed in this study. In

addition, we did not assess specific therapeutic factors, which have been shown to interact with common factors (Boswell, Castonguay, & Wasserman, 2010), might be associated with considerable negative change (Lilienfeld, 2007), and have been shown to foster considerable positive change: Cognitive changes have been shown to precede sudden gains in CBT (Tang, & DeRubeis, 1999; Tang et al., 2005) and therapist interpretation accuracy to predict sudden gains in supportive-expressive therapy (Andrusyna et al., 2006).

One should keep in mind that our associations between the ASC scales and EPD / END do not allow drawing causal inferences. On the one hand, better / worse ASC scale scores could lead to EPD / END. On the other hand, EPD / END could lead to better / worse therapeutic alliance, social support, motivation, and life.

Despite these limitations, our findings provide validity for the ASC social support scale and the ASC life events scale as scales assessing possible problem areas that may account for deterioration in psychotherapy. In clinical practice, when the OQ-Analyst warns therapists that a patient's progress predicts a negative treatment outcome (i. e. when an END occurs) therapists are encouraged to use CSTs, i. e. administer the ASC to patients, attend to items and scale scores that are problematic, use a decision tree to organize problem-solving, and then consider the relevant interventions that are recommended through this approach (Lambert et al., 2007). The results of this study indicate that the ASC social support scale and the ASC life events scale are associated with unusually large negative treatment response.

In contrast to other comparable approaches and methods of outcome informed care that include problem-solving tools (e.g., Bickman, 2005; Duncan, & Miller, 2008; Kopta, & Lowry, 2002), the OQ-Analyst includes measures of social support, motivation and life events rather than relying solely on therapeutic alliance ratings. The results of the present study

provide evidence that the additional variables social support and life events may affect considerable negative change even more than the therapeutic alliance.

Finally, if these results are replicated in other settings, they could suggest a modification of the CSTs of the OQ-Analyst: The current CSTs include a decision-tree with a stepwise procedure: Therapists are encouraged to examine the therapeutic alliance in a first step, to check motivation in a second step, to inspect social support in a third step, and to investigate life events in a fourth step (Lambert et al., 2007). This procedure gives priority to the therapeutic alliance and motivation. CSTs prioritizing variables significantly associated with END - social support and life events in this study - could be more effective in preventing treatment failure.

Conclusions: The quality of life events as well as the quality of social support were associated with considerable positive and negative change, whereas the quality of the therapeutic alliance and the quality of motivation were related only to considerable positive change.

Clinical implications: The results of this study indirectly suggest that encouraging therapists to help deal with life events as well as build a social network can be of importance when trying to improve treatments for poorly responding patients.

References:

Abt, K. (1987). Descriptive data analysis: A concept between confirmatory and exploratory data analysis. *Methods of Information in Medicine*, 26, 77-88.

Aderka, I. M., Nickerson, A., Bøe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *Journal of Consulting & Clinical Psychology*, 80, 80-93.

Amble, I., Gude, T., Stubdal, S., Andersen, B. J., & Wampold, B. E. (2014). The effect of implementing the Outcome Questionnaire-45.2 feedback system in Norway: A multisite randomized clinical trial in a naturalistic setting. *Psychotherapy Research*, epub ahead of print.

Andrusyna, T. P., Luborsky, L., Pham, T., & Tang, T. Z. (2006). The mechanisms of sudden gains in supportive-expressive therapy for depression. *Psychotherapy Research*, 16, 526-536.

Asay, T. P., & Lambert, M. J. (1999). The empirical case for the common factors in therapy: Quantitative findings. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: What works in therapy* (pp. 23-55). Washington, DC: American Psychological Association.

Bankoff, E. A., & Howard, K. I. (1992). The social network of the psychotherapy client and effective psychotherapeutic process. *Journal of Psychotherapy Integration*, 2, 273-294.

Bickman L. (2005). A common factors approach to improving mental health services. *Mental Health Services Research*, 7, 1-4.

Boswell, J. F., Castonguay, L. G., & Wasserman, R. H. (2010). Effects of psychotherapy training and intervention use on session outcome. *Journal of Consulting & Clinical Psychology*, 78, 717-723.

Cavallini, A. Q., & Spangler, D. L. (2013). Sudden gains in cognitive-behavioral therapy for eating disorders. *International Journal of Cognitive Therapy*, 6, 292-310.

Crits-Christoph, P., Ring-Kurtz, S., Hamilton, J. L., Lambert, M. J., Gallop, R., McClure, B. et al. (2012). A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *Journal of Substance Abuse Treatment*, 42, 301-309.

Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive psychotherapy for adult depression: A meta-analysis. *Clinical Psychology Review*, 32, 280-291.

Dinger, U., & Schauenburg, H. (2010). Effects of individual cohesion and patient interpersonal style on outcome in psychodynamically oriented inpatient group psychotherapy. *Psychotherapy Research*, 20, 22-29.

Duncan, B. L., & Miller, S. D. (2008). *The Outcome and Session Rating Scales: The revised administration and scoring manual, including the child outcome rating scale*. Chicago: Institute for the Study of Therapeutic Change.

Finch, A. E., Lambert, M.J., & Schaalje, B.G. (2001). Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. *Clinical Psychology & Psychotherapy*, 8, 231-242.

Greenhalgh, J. (2009). The applications of PROs in clinical practice: What are they, do they work, and why? *Quality of Life Research*, 18, 115-123.

Haas, E., Hill, R. D., Lambert, M. J., & Morrell, B. (2002). Do early responders to psychotherapy maintain treatment gains? *Journal of Clinical Psychology*, 58, 1157-1172.

Hardy, G. E., Cahill, J., Stiles, W. B., Ispan, C., Macaskill, N., & Barkham, M. (2005). Sudden gains in cognitive therapy for depression: A replication and extension. *Journal of Consulting & Clinical Psychology*, 73, 59-67.

Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., et al. (2007). Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. *Psychotherapy Research*, 17, 379-392.

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2012). *Multilevel modeling of categorical outcomes using IBM SPSS*. New York: Routledge.

Hiller, W., Schindler, A., & Lambert, M. J. (2013). Defining response and remission in psychotherapy research: A comparison of the RCI and the method of percent improvement. *Psychotherapy Research*, 22, 1-11.

Hopko, D. R., Robertson, S. M. C., & Carvalho, J. P. (2009). Sudden gains in depressed cancer patients treated with behavioral activation therapy. *Behavior Therapy*, 40, 346-356.

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, 48, 9-16.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, 51, 1059-1064.

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavior therapy of depression. *Clinical Psychology-Science & Practice*, 1, 138-155.

Keller, S. M., Feeny, N. C., & Zoellner, L. A. (2014). Depression sudden gains and transient depression spikes during treatment for PTSD. *Journal of Consulting & Clinical Psychology*, 82, 102-111.

Kopta, S. M., & Lowry, J. L. (2002). Psychometric evaluation of the Behavioral Health Questionnaire-20: A brief instrument for assessing global mental health and the three phases of psychotherapy outcome. *Psychotherapy Research*, 12, 413-426.

Lambert, M. J. (2005). Early response in psychotherapy: Further evidence for the importance of common factors rather than "placebo effects". *Journal of Clinical Psychology*, 61, 855-869.

Lambert, M. J. (2012). Helping clinicians to use and learn from research-based systems: The OQ-Analyst. *Psychotherapy*, 49, 109-114.

Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, 50, 42-51.

Lambert, M. J., Bailey, R., Kimball, K., Shimokawa, K., Harmon, S. C., & Slade, K. (2007). *Clinical support tools manual-brief version-40*. Salt Lake City: OQMeasures.

Lambert, M. J., Hannöver, W., Nisslmüller, K., Richard, M., & Kordy, H. (2002). Questionnaire on the results of psychotherapy: Reliability and validity of the German translation of the Outcome Questionnaire 45.2 (OQ-45.2). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 31, 40-47.

Lambert, M. J., Kahler, M., Harmon, S. C., Burlingame, G. M., & Shimokawa, K. (2011). *Administration and scoring manual for the Outcome Questionnaire-45.2*. Salt Lake City: OQMeasures.

Lambert, M. J., Whipple, J.L., Bishop, M. J., Vermeersch, D. A., Gray, G. V., & Finch, A. E. (2002). Comparison of empirically derived and rationally derived methods for identifying clients at risk for treatment failure. *Clinical Psychology & Psychotherapy*, 9, 149-164.

Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science*, 2, 53-70

Lutz, W., Bachmann, F., Tschitsaz, A., Smart, D. W., & Lambert, M. J. (2007). Temporal and sequential patterns of nonlinear change psychotherapy. The phenomenon of sudden gains and sudden losses in its context. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 36, 261-269.

Lutz, W., Ehrlich, T., Rubel, J., Hallwachs, N., Röttger, M. A., Jorasz, C. et al. (2013). The ups and downs of psychotherapy: Sudden gains and sudden losses identified with session reports. *Psychotherapy Research*, 23, 14-24.

Mander, J., Schlarb, A., Teufel, M., Keller, F., Hautzinger, M., Zipfel, S., Sammet, I. (2014). The Individual Therapy Process Questionnaire: Development and validation of a revised measure to evaluate general change mechanisms in psychotherapy. *Clinical Psychology & Psychotherapy*, published online.

Mander, J. V., Wittorf, A., Schlarb, A., Hautzinger, M., Zipfel, S., & Sammet, I. (2013). Change mechanisms in psychotherapy: Multiperspective assessment and relation to outcome. *Psychotherapy Research*, 23, 105–116.

McCarthy, K. S., & Barber, J. P. (2009). The Multitheoretical List of Therapeutic Interventions (MULTI): Initial report. *Psychotherapy Research*, 19, 96-113.

Newnham, E. A., Hooke, G. R., & Page, A. R. (2010). Progress monitoring and feedback in psychiatric care reduces depressive symptoms. *Journal of Affective Disorders*, 127, 139-146.

Pelletier, L. G., Tuson, K. M., & Haddad, N. K. (1997). Client Motivation for Therapy Scale: A measure of intrinsic motivation, extrinsic motivation, and amotivation for therapy. *Journal of Personality Assessment*, 68, 414-435.

Probst, T., Lambert, M. J., Loew, T. H., Dahlbender, R. W., Göllner, R., & Tritt, K. (2013). Feedback on patient progress and clinical support tools for therapists: Improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice. *Journal of Psychosomatic Research*, 75, 255-261.

Probst, T., Lambert, M. J., Dahlbender, R. W., Loew, T. H., & Tritt, K. (2014) Providing feedback on patient progress and clinical support tools for therapists: Is the therapeutic process of patients on-track to recovery improved in psychosomatic in-patient therapy under the conditions of routine practice? *Journal of Psychosomatic Research*, 76, 477-484.

Sexton, H. (1996). Process, life events, and symptomatic change in brief eclectic psychotherapy. *Journal of Consulting & Clinical Psychology*, 64, 1358-1365.

Simon, W., Lambert, M. J., Harris, M. W., Busath, G., & Vazquez, A. (2012). Providing patient progress information and clinical support tools to therapists: Effects on patients at risk of treatment failure. *Psychotherapy Research*, 22, 638-647.

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a

psychotherapy quality assurance system. *Journal of Consulting & Clinical Psychology*, 78, 298-311.

Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology & Psychotherapy*, 15, 287-303.

Spielmans, G. I., Masters, K. S., & Lambert, M. J. (2006). A comparison of rational versus empirical methods in prediction of negative psychotherapy outcome. *Clinical Psychology & Psychotherapy*, 13, 202-214.

Tang, T. Z., & DeRubeis, R.J. (1999) Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting & Clinical Psychology*, 67, 894-904.

Tang, T. Z., DeRubeis, R. J., Beberman, R., & Pham, T. (2005). Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. *Journal of Consulting & Clinical Psychology*, 73, 168-172.

Tang, T. Z., Luborsky, L., & Andrusyna, T. (2002). Sudden gains in recovering from depression: Are they also found in psychotherapies other than cognitive-behavioral therapy? *Journal of Consulting & Clinical Psychology*, 70, 444-447.

Tschitsaz-Stucki, A., & Lutz, W. (2009). Identification and clarification of sudden gains and sudden losses in individual psychotherapy courses. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 38, 13-23.

Whipple, J. L., Lambert, M. J., Vermeersch, D.A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology, 50*, 59-68.

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., Erikson, P., & ISPOR Task Force for Translation and Cultural Adaptation (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health, 8*, 94-104.

World Health Organization WHO (1992). *ICD-10: The ICD-10 classification of mental and behavioural disorders*. Geneva, WHO.

Tables:

Table 1: Sample characteristics.

Gender n (%)	Female	161 (59.4)
	Male	110 (40.6)
Education n (%)	Higher secondary school	53 (19.6)
	Intermediate secondary school	76 (28.0)
	Lower secondary school	118 (43.5)
	Other type of graduation	24 (8.9)
Amount of patients with specific diagnoses n (%)	Depressive disorders (F32 / F33)	178 (65.7)
	Anxiety disorders (F40 / F41)	69 (25.5)
	Obsessive compulsive disorders (F42)	8 (3.0)
	Reaction to severe stress, and adjustment disorders (F43)	42 (15.5)
	Somatoform disorders (F45)	158 (58.3)
	Eating disorders (F50)	35 (12.9)
	Disorders of adult personality and behavior (F6)	20 (7.4)
	Other disorders (all F diagnoses except the already mentioned)	57 (21.0)
Number of diagnoses according to chapter V of the ICD-10 M (SD)		2.26 (1.06)
Age M (SD)		48.36 (13.08)

OQ total score of the intake week M (SD)	81.52 (25.70)
Treatment weeks from intake to discharge M (SD)	4.65 (2.13)

Table 2: Results of the multilevel model with the extreme positive deviation classification (EPD) as binary outcome (treatment week with EPD=1; treatment week without EPD=0) and the scales of the “Assessment for Signal Cases” (ASC) as time-dependent covariates.

Parameter	Regression coefficient β	Standard error	95% Wald confidence interval		Wald-chi-square	Degrees of freedom	p-value	Exp(β)
			Lower bound	Upper bound				
Intercept	-1.640	0.151	-1.936	-1.344	117.863	1	<0.01	0.194
ASC therapeutic alliance scale	0.052	0.022	0.009	0.094	5.732	1	0.017	1.053
ASC social support scale	0.066	0.014	0.039	0.094	22.482	1	<0.01	1.069
ASC motivation scale	0.093	0.032	0.031	0.155	8.595	1	<0.01	1.097
ASC life events scale	0.081	0.020	0.043	0.120	17.053	1	<0.01	1.085

Table 3: Results of the multilevel model with the extreme negative deviation classification (END) as binary outcome (treatment week with END=1; treatment week without END=0) and the scales of the “Assessment for Signal Cases” (ASC) as time-dependent covariates.

Parameter	Regression coefficient β	Standard error	95% Wald confidence interval		Wald- chi-square	Degrees of freedom	p-value	Exp(β)
			Lower bound	Upper bound				
Intercept	-2.865	0.265	-3.385	-2.345	116.513	1	<0.01	0.057
ASC therapeutic alliance scale	-0.009	0.026	-0.060	0.041	0.136	1	0.713	0.991
ASC social support scale	-0.077	0.022	-0.121	-0.034	12.127	1	<0.01	0.926
ASC motivation scale	-0.044	0.033	-0.109	0.021	1.727	1	0.189	0.957
ASC life events scale	-0.087	0.025	-0.135	-0.039	12.441	1	<0.01	0.917

IV. Diskussion und Ausblick

In diesem Kapitel werden die Ergebnisse der Arbeit integriert und daraus mögliche Fragestellungen für zukünftige Forschungsarbeiten abgeleitet. Abschließend erfolgt eine zusammenfassende Betrachtung.

IV. 1. Publikationen 1 & 2

In zwei Publikationen wurde die Feedbackintervention OQ-Analyst in der stationären Psychosomatik mit TAU verglichen (Probst et al., 2013; Probst, Lambert et al., 2014). Die Feedbackintervention bestand aus Feedback über den Therapieverlauf und CST an den verantwortlichen Einzeltherapeuten. Während durch diese Intervention der globale Outcome von NOT Patienten bei Behandlungsende verbessert werden konnte (Probst et al., 2013), optimierte diese Art von Feedback bei OT Patienten den Therapieprozess nur kurzfristig für eine Woche und nur geringfügig bzgl. globaler Symptombelastung (Probst, Lambert et al., 2014).

Das Ergebnis, dass Feedbackinterventionen mehr bei NOT Patienten als bei OT Patienten zur Therapieoptimierung beitragen, wurde auch in anderen Studien in anderen Settings gefunden (siehe z. B. Lambert, 2010; Shimokawa et al., 2010).

Die Ergebnisse von Probst et al. (2013) können als weiterer Beitrag zur Evidenzbasierung der Wirksamkeit von Feedbackinterventionen zur Prävention von Verschlechterungen gewertet werden. Es liegen nun insgesamt vier „Randomized Clinical Trials“ (RCT; Lambert et al., 2001; Probst et al., 2013; Simon et al., 2012; Whipple et al., 2003) und zwei Meta-Analysen (Lambert et al., 2003; Shimokawa et al., 2010) vor, die den positiven Einfluss des Feedbacksystems OQ-Analyst auf den Outcome von NOT Patienten bei Behandlungsende zeigen konnten. Weitere RCT zu anderen Feedbacksystemen (z. B. Reese et al., 2009; Anker, Duncan, & Sparks, 2009) sowie Meta-Analysen über verschiedene Feedbacksysteme (z. B. Knaup, Kösters, Schöfer, Becker, & Puschner, 2009) bestätigten den positiven Einfluss von Feedback auf die Ergebnisqualität von Psychotherapien. Diesen beiden Studientypen (RCT und Meta-Analysen) werden in der EBM hohe Evidenzgrade zugeschrieben (Harbour, & Miller, 2001).

Die vorliegende wissenschaftliche Evidenz zum Nutzen von Feedback zur Reduktion von Verschlechterungsraten sowie die gesetzliche Forderung an Leistungserbringer im Gesundheitssystem sich an Maßnahmen zur Steigerung der Ergebnisqualität zu beteiligen (Sozialgesetzbuch V, §135a), könnten eine Implementierung von Monitoring und Feedback in die

psychotherapeutische Versorgungspraxis nahe legen. Jedoch wäre zur effektiven Implementierung zuerst eine Veränderung der stark negativ gefärbten Einstellung von Therapeuten gegenüber Monitoring und Feedback (Aoun, Pennebaker, & Janca, 2002; Jensen-Doss, & Hawley, 2010; Walter, Cleary, & Rey, 1998) vonnöten, z. B. durch finanzielle Anreize (Lilienfeld et al., 2013). Als möglicher Grund für diese negative Einstellung sowie als ein Argument gegen eine Implementierung in die Routinepraxis könnte angeführt werden, dass Therapeuten durch ROM sowie durch Feedback zusätzlichen zeitlichen Aufwand haben, der in der Routineversorgung nicht zur Verfügung steht. Jedoch kann dieser Aufwand aufgrund technischer Möglichkeiten gering gehalten werden, z. B. indem Patienten regelmäßig über Computer (Tablets, Pocket-PCs) Fragebögen bearbeiten und die weiteren Schritte (Fragebogenauswertung, Erstellen von Feedbackberichten, Senden der Feedbackberichte an die Therapeuten) automatisiert erfolgen.

Zur weiteren Evidenzbasierung von Feedbackinterventionen sind in zukünftigen Studien v. a. follow-up Untersuchungen nötig, um die längerfristigen Effekte von Feedback über das Behandlungsende hinaus zu untersuchen. Bisher hat nur eine Studie in der Psychiatrie³ (Byrne, Hooke, Newnham, & Page, 2012) untersucht, wie sich Feedback während der Behandlung auf den längerfristigen Verlauf nach der Behandlung auswirkt: OT Patienten zeigten in einem 6-Monats-follow-up signifikant weniger Klinikaufenthalte, wenn ihre Therapeuten während des Psychiatrieaufenthaltes Feedback über den Therapieverlauf erhalten haben. Dieser längerfristig positive Effekt von Feedback zeigte sich jedoch nicht für NOT Patienten (Byrne et al., 2012), für die dieser Effekt eher zu erwarten gewesen wäre als für OT Patienten, denn zum Zeitpunkt Behandlungsende war der Outcome von NOT Patienten aber nicht der von OT Patienten durch Feedback erhöht (Newnham, Hooke, & Page, 2010). Byrne et al. (2012) bezifferten den finanziellen Gewinn durch die geringeren Klinikaufenthalte der OT Patienten im 6-Monats-follow-up auf AUD \$ 60.000 pro 100 Patienten. Im Vergleich dazu erscheinen die finanziellen Kosten eines Monitoring- und Feedbacksystems (z. B. OQ-Analyst: USD \$ 200 pro Kliniker pro Jahr; Lambert, 2012) gering. Da der Studie von Byrne et al. (2012) ein quasiexperimentelles Studiendesign zugrunde lag, könnten die berichteten Ergebnisse aber auch von Störvariablen (mit-) beeinflusst worden sein. Zukünftige Studien zu

³ Im Gegensatz zu dieser Arbeit und zu den anderen Feedbackstudien, in denen die OQ-45 Algorithmen eingesetzt wurden (siehe II.), basierte der Therapieverlauf in jenem Psychiatrieprojekt (Newnham et al., 2010; Byrne et al., 2012) zur Klassifikation von OT / NOT Patienten nicht auf einem Beschwerdefragebogen, sondern auf einem Fragebogen zum Wohlbefinden.

längerfristigen Feedbackeffekten sollten daher im randomisierten Design durchgeführt werden.

Die kurz- bzw. langfristige Effektivität von Feedbackinterventionen könnte zudem von Therapeutenvariablen abhängig sein, da die Effektivität psychotherapeutischer Interventionen u. a. auch vom Therapeuten abhängt (Okiishi, Lambert, Nielsen, & Ogles, 2003). In zukünftigen Feedbackstudien sollte daher die Therapeutenvariable mitberücksichtigt werden, da dies im hier berichteten Projekt aufgrund der Absprachen mit den Kliniken und den Behandlungsteams nicht möglich war.

Im Rahmen der „Contextual Feedback Intervention Theory“ (CFIT; Sapyta, Riemer, & Bickman, 2005) wird postuliert, dass bestimmte Therapeuteneigenschaften („goal commitment“, „cognitive dissonance“, „personal responsibility“) mit der Feedbackeffektivität - im Sinne einer Verbesserung der Behandlungsqualität - zusammenhängen dürften. In einer empirischen Studie fanden de Jong, van Sluis, Nugter, Heisser und Spinhoven (2012), dass die Therapeuteneigenschaften Selbstwirksamkeit und Commitment bzgl. Gebrauch von Feedback einen signifikanten Einfluss auf die Feedbackeffektivität hatten. Simon et al. (2012) haben keine Therapeuteneigenschaften erhoben, aber Feedbackeffekte getrennt nach Therapeuten analysiert. Sie fanden, dass der Effekt von Feedback über den Therapieverlauf und CST auf NOT Patienten bei manchen Therapeuten deutlicher ist als bei anderen Therapeuten. Anker et al. (2009) berichten, dass Feedback besonders bei solchen Therapeuten effektiv sei, die ohne Feedback schlechte Outcomes erzielen. Durch mehr Forschung zum Zusammenhang zwischen Therapeuteneigenschaften einerseits und Feedbackeffektivität andererseits, könnten Therapeuten identifiziert werden, bei denen sich der Einsatz von Feedback zur Optimierung ihrer Behandlungsqualität besonders lohnt.

Die Effektivität von Feedback hängt nach der CFIT auch mit spezifischen Elementen der Feedbackberichte („content“, „sign“, „format“) zusammen. Bisher liegt zur Fragestellung der Nützlichkeit verschiedener Feedbackelemente eine deskriptive Arbeit von Haderlie (2009) zum OQ-Analyst vor. Therapeuten bewerteten folgende Elemente der Feedbackberichte des OQ-Analyst als besonders hilfreich: Grafische Darstellung des Therapieverlaufs, Kategorisierung der Veränderung seit Behandlungsbeginn und Angaben zum aktuellen Belastungsgrad. Einem aktuellen Review zufolge (Boyce, Browne, & Greenhalgh, 2014) erhöht leichte Interpretierbarkeit der Feedbackberichte und hoher Nutzen für die klinische Entscheidungsfindung die Wahrscheinlichkeit, dass Feedback effektiv genutzt wird. Es fehlen hier jedoch bisher

empirische Studien, die untersucht haben, ob unterschiedlich gestaltete Feedbackberichte einen unterschiedlichen Effekt auf das Behandlungsergebnis von Patienten haben.

Bzgl. Gestaltung der Feedbackberichte könnte zukünftig z. B. untersucht werden, ob Feedback unterschiedlich effektiv ist, je nachdem welches behandlungsrelevante Kriterium in den Feedbackberichten rückgemeldet wird. Man könnte spekulieren, dass die Effektivität von Feedback davon abhängt, in welcher Therapiephase welches behandlungsrelevante Kriterium rückgemeldet wird: Wie unter II. 1. beschrieben, steht in der Remoralisierungsphase die Verbesserung des Wohlbefindens im Vordergrund, entsprechend könnte hier Feedback zum Wohlbefinden besonders effektiv sein. In der Remediationsphase dagegen, wenn es um Symptomreduktion geht, könnte Feedback über Symptome positivere Auswirkungen haben als Feedback zum Wohlbefinden. In der Rehabilitationsphase schließlich könnte zur Verbesserung der interpersonellen und sozialen Probleme v. a. Feedback zu interpersonellen und sozialen Problemen nützlich sein.

Bezogen auf die Gestaltung von Feedbackberichten könnte auch evaluiert werden, ob CST unterschiedlich effektiv sind, je nachdem welche Elemente im Rahmen der CST rückgemeldet werden. Die CST des OQ-Analyst beinhalten zwei Komponenten: Erstens Feedback über mögliche Einflussfaktoren auf den Therapieverlauf und zweitens einen „Decision Tree“, der zum therapeutisch adäquaten Umgang mit denjenigen der möglichen Einflussfaktoren anleiten soll, die beim Patienten problematisch ausgeprägt sind. In den bisherigen Studien wurden beide Elemente immer zusammen verwendet. Es ist daher unklar, welchen Mehrwert der „Decision Tree“ zusätzlich zum Feedback über die möglichen Einflussfaktoren hat. Dies könnte empirisch ermittelt werden, z. B. indem die Behandlungsqualität von Therapeuten, die Feedback über die möglichen Einflussfaktoren ohne den „Decision Tree“ erhalten, verglichen wird mit der Behandlungsqualität von Therapeuten, die Feedback über die möglichen Einflussfaktoren plus den „Decision Tree“ bekommen. Wie unter II. 2. erwähnt, wurden die CST des OQ-Analyst für NOT Patienten entwickelt und integrieren Aspekte, die sich in früheren Studien als einflussreich auf den psychotherapeutischen Outcome / Prozess gezeigt haben. CST für NOT Patienten könnten noch effektiver sein, wenn sie einen Schwerpunkt auf diejenigen Faktoren legen, die mit NOT Klassifikationen assoziiert sind. Erste Hinweise darauf, welche Faktoren empirisch mit NOT Klassifikationen zusammenhängen, liefert die Arbeit von Probst et al. (under re-review), auf die weiter unten eingegangen wird. In dieser Dissertation wurden die für NOT Patienten entwickelten CST des OQ-Analyst zum ersten Mal auch für OT Patienten verwendet, was jedoch TAU nicht wesentlich verbessern konnte. OT Patienten könnten evtl. stärker profitieren, wenn Therapeuten CST erhalten, die an OT Patienten

angepasst bzw. speziell für diese entwickelt sind (z. B. Feedback nicht nur über negativ ausgeprägte Einflussfaktoren wie im OQ-Analyst, siehe Abbildung 2, sondern auch Feedback über positive Ergebnisse).

Ein bisher in der Feedbackforschung vernachlässigter Bereich betrifft die Supervision. Feedbackinterventionen könnten an Effektivität gewinnen, wenn Therapeuten angeregt werden, die Feedbackberichte mit ihrem Supervisor zu diskutieren. Bisher wurden nur bei Reese et al. (2009) Therapeuten motiviert, die Feedbackberichte mit Supervisoren zu besprechen. Im Ausbildungskontext wurde diese Feedbackintervention mit TAU ohne Feedback verglichen mit dem Ergebnis einer besseren Behandlungsqualität der Feedbackgruppe. Um den zusätzlichen Wert von Feedback an den Supervisor zu evaluieren, würde sich ein Vergleich zwischen Feedback an Supervisand plus Supervisor vs. Feedback nur an den Supervisand anbieten. Solch ein Studiendesign würde jedoch auch das Problem mit sich bringen, dass Supervisanden dann nicht mehr frei darüber entscheiden dürften, was sie von ihren Patientenbehandlungen mit dem Supervisor besprechen und was nicht.

Überdies könnte man untersuchen, ob die Qualität der Supervision zunimmt, wenn Supervisoren Feedback über ihre Supervisionsqualität aus Supervisandensicht erhalten. Da Supervision eine Verbesserung der beruflichen Kompetenz anstrebt, könnte eine Verbesserung der Supervision durch Feedback auch dazu führen, dass Supervisanden effektivere Patientenbehandlungen durchführen. Wie sich Feedback von Supervisand an Supervisor auf die Qualität der Supervisionssitzungen bzw. auf die Behandlungsqualität der Supervisanden auswirkt, könnte in zukünftigen Forschungsprojekten untersucht werden.

Eine Erhöhung der psychotherapeutischen Kompetenz könnte durch Feedback auch erreicht werden, indem man Therapeuten kontinuierlich Feedback darüber gibt, wie gut sie bestimmte psychotherapeutische Kompetenzen in ihren Sitzungen erfüllen (Kaufmann, & Weck, 2014). Hierzu könnten zukünftige Arbeiten prüfen, ob Therapeuten, die solch ein Kompetenzfeedback erhalten, effektivere Patientenbehandlungen durchführen.

IV. 2. Publikation 3

Neben den Fragestellungen zum Einfluss einer Feedbackintervention auf den Therapieoutcome bzw. Therapieprozess wurden in dieser Dissertation auch Eingangsbelastungen und wöchentliche Änderungsmuster im OQ-45 und ISR untersucht (Probst, Loew et al., 2014): Patienten schätzten sich während der Aufnahmewoche in den erfassten behandlungsrelevanten

Kriterien als unterschiedlich belastet ein (am stärksten bzgl. depressiver Symptombelastung, am geringsten hinsichtlich Essstörungssymptomatik) und zeigten in diesen Kriterien differenzielle Therapieverläufe in den ersten vier Wochen stationärer psychosomatischer Therapie: Belastungen durch depressive Symptome sowie durch Angstsymptome besserten sich vor Somatisierungsbeschwerden, während sich Belastungen durch Zwangs- sowie Essstörungssymptome während der vier untersuchten Behandlungswochen noch nicht signifikant verbesserten. Insgesamt ging die Gesamtbelastung durch psychische Beschwerden zu Beginn der stationären psychosomatischen Behandlung deutlicher zurück als in den darauf folgenden Wochen. Dieser Befund hat Ähnlichkeit zur negativ beschleunigten Änderungskurve, wie sie in der ambulanten Psychotherapie mit session-by-session Analysen gefunden wurde (z. B. Howard, Kopta, Krause, & Orlinsky, 1986; Stulz, Lutz, Kopta, Minami, & Saunders, 2013). Im stationären Setting könnten die deutlicheren Verbesserungen zu Beginn der Behandlung besonders mit dem „common factor“ „healing environment“ (Frank, & Frank, 1991) zusammenhängen.

Dass sich in stationärer psychosomatischer Behandlung unterschiedliche spezifische Symptombelastungen differenziell verändern, könnte dafür sprechen, in Feedbackberichten die Therapieverläufe mehrerer spezifischer Symptombelastungen rückzumelden anstatt (wie im OQ-Analyst) einen Therapieverlauf zur psychischen Gesamtbelastung. Z. B. könnte für einen Patienten sein individueller Therapieverlauf bzgl. depressiver Symptombelastung verglichen werden mit dem erwarteten Therapieverlauf der depressiven Symptombelastung, sein individueller Therapieverlauf bzgl. zwangsbezogener Symptombelastung mit dem erwarteten Therapieverlauf der zwangsbezogenen Symptombelastung etc.. In Zukunft könnte untersucht werden, ob solche Feedbackberichte, die mehrere Therapieverläufe zu mehreren spezifischen Symptombelastungen beinhalten, Vor- oder Nachteile mit sich bringen (vgl. die obigen Ausführungen zur CFIT und zur Gestaltung von Feedbackberichten).

In zukünftigen Arbeiten zu Monitoring und Feedback in der stationären Psychosomatik sollten erwartete Therapieverläufe bestimmt werden, die spezifisch für das Setting der stationären Psychosomatik sind. Die in dieser Dissertation verwendeten empirisch ermittelten erwarteten Therapieverläufe des OQ-45 basierten auf verschiedenen psychotherapeutischen Settings der Vereinigten Staaten von Amerika. Durch settingsspezifische erwartete Therapieverläufe könnten NOT und OT Patienten präziser erfasst werden.

Settingsspezifische Therapieverläufe für die stationäre Psychosomatik sollten „empirically derived expected recovery curves“ sein, da empirisch ermittelte erwartete Therapiever-

läufe den erwarteten Therapieverläufen, wie sie Experten einschätzen („rationally derived expected recovery curves“), hinsichtlich der Vorhersage von Verschlechterungen überlegen sind (Lambert, Whipple, Bishop et al., 2002; Lutz et al., 2006; Spielmans et al., 2006).

Eine der zentralen Fragen bei der Generierung von erwarteten Therapieverläufen ist, welche Patientenmerkmale bei Behandlungsbeginn als Basis für den erwarteten Therapieverlauf eines Patienten dienen sollen. Bei den erwarteten Therapieverläufen des OQ-45 war dies jeweils einzig der Eingangswert im OQ-45. Obwohl in den Untersuchungen zum OQ-45 nach Lambert (2013) der Einbezug zusätzlicher Variablen kaum zur Erhöhung der Güte der erwarteten Therapieverläufe beitrug, könnten in der stationären Psychosomatik auch andere Variablen relevant sein: Z. B. waren in früheren Studien zur Psychosomatik soziodemographische Variablen (z. B. beruflicher Status, Berentung), das Bindungsmuster, die Diagnosegruppe, die psychische und somatische Komorbidität, der Pessimismus und die Selbstwirksamkeit signifikante Prädiktoren des Outcomes (z. B. Fliege, Rose, Bronner, & Klapp, 2002; Herrmann, & Huber, 2013; Huber, Albrecht, Hautum, Henrich, & Klug, 2009; Strauss et al., 2006). Lutz, Martinovich und Howard (1999) berichteten in diesem Zusammenhang, dass *zusätzlich zum Fragebogen-Eingangswert* auch die Therapieerwartung sowie die Chronizität der Störung signifikant mit dem Therapieverlauf im ambulanten Setting zusammenhängen. Sollten sich in zukünftigen Studien zur stationären Psychosomatik mehrere Variablen als relevante Prädiktoren des Therapieverlaufs herausstellen, so wäre der parallele Einbezug dieser Variablen in die Generierung von erwarteten Therapieverläufen mittels der „nearest neighbors“ Methode (Lutz et al., 2005) möglich. Bei dieser Methode werden zur Generierung des Therapieverlaufs eines bestimmten Patienten X diejenigen Patienten aus einer Normstichprobe ausgewählt, die dem Patienten X über alle der ausgewählten Variablen (z. B. Beschwerdeausmaß, Therapieerwartung, Chronizität) am ähnlichsten sind.

Die Arbeit von Probst, Loew et al. (2014) liefert auch Hinweise dafür, dass die OQ-45 Skalen und die meisten der ISR Skalen (außer Zwangs- und Essstörungsskala) in den ersten vier Behandlungswochen stationärer psychosomatischer Therapie das testpsychologische Nebenkriterium Änderungssensitivität erfüllen (zur Änderungssensitivität des OQ-45 siehe auch Haug, Puschner, Lambert, & Kordy, 2004) und damit zum ROM in der stationären Psychosomatik geeignet sind.

Die fehlenden signifikanten Veränderungen in der ISR Zwangs- und Essstörungsskala sollten jedoch nicht als Beweis für eine geringe Änderungssensitivität dieser ISR Skalen gewertet werden: Auch andere Einflüsse könnten hier eine Rolle spielen: Z. B. zeigten die Pati-

enten in der ISR Zwangs- und Essstörungsskala während der Aufnahmewoche geringe Belastungsgrade, so dass relativ wenig Spielraum für Verbesserungen in diesen Skalen bestand. Zudem könnten die zwei beteiligten Kliniken einen therapeutischen Schwerpunkt haben, der nicht auf die Veränderung von Zwangs- bzw. Essstörungssymptomen, sondern auf die Verbesserung anderer Symptombereiche ausgerichtet ist. Weiterhin könnten sich bei Symptomen einer Zwangs- bzw. Essstörung erst im weiteren Therapieverlauf deutlichere Veränderungen einstellen als in den ersten vier Behandlungswochen, da diese zwei Störungsbilder in der Regel mit einer längeren Behandlungsdauer einhergehen (von Heymann et al., 2003). Ab welcher Behandlungswoche sich Zwangs- bzw. Essstörungssymptome in der stationären Psychosomatik verbessern, könnte Gegenstand zukünftiger Studien sein.

IV. 3. Publikation 4

Mit den wöchentlichen Monitoringdaten des OQ-45 und des ASC untersuchten Probst et al. (under re-review), ob bedeutsame Verbesserungen bzw. Verschlechterungen mit den vier Skalen des ASC zusammenhängen: Therapiebeziehung, soziale Unterstützung, Motivation und Lebensereignisse. Als bedeutsame Verbesserungen wurden diejenigen der wöchentlichen Messungen definiert, bei denen im OQ-Analyst ein „Blue“ Signal auftrat (siehe II. 2.). Diejenigen der wöchentlichen Messungen, in denen der OQ-Analyst ein „Yellow“ oder „Red“ Signal anzeigte (NOT Klassifikationen), wurden als bedeutsame Verschlechterungen definiert (siehe II. 2.). Hinsichtlich der Untersuchung des OQ-Analyst waren hier besonders die Zusammenhänge zwischen den ASC Skalen und bedeutsamen Verschlechterungen von Interesse, denn beim Einsatz des OQ-Analyst in der klinischen Routineversorgung sollen Therapeuten diejenigen Patienten, die eine bedeutsame Verschlechterung (d. h. eine NOT Klassifikation) zeigen, den ASC ausfüllen lassen, Feedback über die ASC Ergebnisse erhalten und in einem „Decision Tree“ zum therapeutisch adäquaten Umgang mit problematisch ausgeprägten ASC Ergebnissen angeleitet werden (siehe II. 2.). Diesem Vorgehen liegt die Annahme zugrunde, dass die im ASC erfassten Faktoren mit bedeutsamer Verschlechterung zusammenhängen und eine Berücksichtigung sowie ein adäquater Umgang mit diesen Faktoren Verschlechterungsraten reduzieren könnte. Zu der Annahme, dass die im ASC erhobenen Faktoren tatsächlich mit bedeutsamer Verschlechterung assoziiert sind, lagen bisher keine empirischen Studien vor.

Probst et al. (under re-review) zeigten, dass alle vier Faktoren des ASC mit bedeutsamen Verbesserungen zusammenhängen, jedoch nur zwei der ASC Faktoren - soziale Unterstützung und Lebensereignisse - mit bedeutsamen Verschlechterungen assoziiert waren.

Sollten auch zukünftige Studien zeigen, dass die Qualität der Lebensereignisse sowie der sozialen Unterstützung stärker mit bedeutsamen Verschlechterungen assoziiert ist als die Qualität der Therapiebeziehung und der Motivation, könnte dies eine Veränderung der CST im OQ-Analyst nahe legen: Bei Patienten mit bedeutsamer Verschlechterung werden die Therapeuten im derzeitigen „Decision Tree“ des OQ-Analyst dazu angeleitet, zuerst die Therapiebeziehung und die Motivation zu prüfen, bevor sie ihre Aufmerksamkeit auf die soziale Unterstützung und die Lebensereignisse richten (siehe II. 2.). Dieser Ablauf erweckt den Eindruck, dass der Therapiebeziehung und der Motivation Priorität eingeräumt wird. CST, die der sozialen Unterstützung und den Lebensereignissen ebenso große oder sogar größere Bedeutung beimessen als der Therapiebeziehung und der Motivation könnten noch stärker zur Prävention von Verschlechterungen in der stationären Psychosomatik beitragen als die bisherigen CST des OQ-Analyst.

In der Arbeit von Probst et al. (under re-review) wurden die Faktoren Therapiebeziehung, soziale Unterstützung, Motivation und Lebensereignisse jeweils global erfasst. Diese Faktoren lassen sich jedoch noch weiter differenzieren: Emotionale Bindung, Übereinstimmung bzgl. Therapieziele und Übereinstimmung hinsichtlich therapeutischer Aufgaben werden als unterschiedliche Facetten der Therapiebeziehung angesehen (Wilmers et al., 2008). Auch bei der sozialen Unterstützung wird z. B. zwischen emotionaler Unterstützung, praktischer Unterstützung, sozialer Integration und sozialer Belastung differenziert (Fydrich, Geyer, Hessel, Sommer, & Brähler, 1999). Und auch die Motivation kann untergliedert werden z. B. in Behandlungserwartungen, Krankheitserleben, Laienätiologie und Offenheit für Psychotherapie (Schneider, Klauer, Janssen, & Tezlaff, 1999). Zukünftige Studien könnten der Fragestellung nachgehen, wie stark spezifische Facetten der in dieser Arbeit global erfassten Faktoren mit bedeutsamen Veränderungen zusammenhängen: Es könnte z. B. sein, dass - selbst wenn die globale Therapiebeziehung nicht mit bedeutsamen Verschlechterungen zusammenhängt - spezifische Facetten der Therapiebeziehung sehr wohl mit bedeutsamen Verschlechterungen assoziiert sind. Auch dies sollte bei den oben diskutierten Änderungsmöglichkeiten der CST des OQ-Analyst bedacht werden.

Eine Einschränkung bei der Studie von Probst et al. (under re-review) ist, dass keine therapieschulenspezifischen Faktoren (z. B. Kognitive Umstrukturierung, Konfrontation) erhoben wurden, diese aber eine Rolle bei bedeutsamen Veränderungen spielen (Andrusyna, Luborsky, Pham, & Tang, 2006; Tang, & DeRubeis, 1999; Tang, DeRubeis, Beberman, & Pham, 2005; Tschitsaz-Stucki, & Lutz, 2009). Zukünftige Forschungsarbeiten könnten daher untersuchen, wie Interaktionen zwischen therapieschulenspezifischen Faktoren und den im ASC erhobenen Faktoren mit bedeutsamen Veränderungen assoziiert sind. Bisher wurden Interaktionen zwischen therapieschulenspezifischen Faktoren und therapieschulenübergreifenden Faktoren („common factors“, z. B. Therapiebeziehung) lediglich bezogen auf den Sitzungsergebnis analysiert (Boswell, Castonguay, & Wasserman, 2010): Patienten schätzten ihre therapeutischen Sitzungen als weniger hilfreich ein, wenn Therapeuten überdurchschnittlich viel „common factors“ und gleichzeitig therapieschulenspezifische Faktoren der kognitiven Verhaltenstherapie in den Sitzungen einsetzten.

IV. 4. Zusammenfassende Betrachtung

In der Gesamtschau betrachtet, zeigt diese Arbeit folgende Ergebnisse, Stärken und Schwächen. Als erstes Ergebnis kann festgehalten werden, dass die Feedbackintervention OQ-Analyst in der stationären Psychosomatik zwar zur Prävention von Verschlechterungsraten, aber nur bedingt zur Optimierung unproblematischer Therapieverläufe beitragen kann. Ein weiteres Ergebnis dieser Arbeit ist, dass sich unterschiedliche Symptombelastungen während der ersten vier Behandlungswochen differenziell verändern und der OQ-45 sowie das ISR zum ROM in der stationären Psychosomatik geeignet sind. Als weiteres Ergebnis zeigte sich, dass die Qualität der sozialen Unterstützung sowie der Lebensereignisse sowohl mit bedeutsamen Verbesserungen als auch mit bedeutsamen Verschlechterungen assoziiert ist, die Qualität der therapeutischen Beziehung und der Motivation hingegen nur mit bedeutsamen Verbesserungen. Als Stärken der Arbeit können die Durchführung unter den Bedingungen der realen Versorgungspraxis in Kliniken der Routineversorgung (externe Validität) sowie das randomisierte Design zur Evaluation der Feedbackintervention OQ-Analyst (interne Validität) angeführt werden. Schwächen der Arbeit sind die relativ hohe drop-out Rate und die Verwendung von erwarteten Therapieverläufen, die auf verschiedenen psychotherapeutischen Settings der Vereinigten Staaten von Amerika und nicht auf dem stationären psychosomatischen Setting Deutschlands basieren.

V. Literaturverzeichnis zum Klappentext

Aderka, I. M., Nickerson, A., Bøe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 80, 80-93.

Amble, I., Gude, T., Stubdal, S., Andersen, B. J., & Wampold, B. E. (2014). The effect of implementing the Outcome Questionnaire-45.2 feedback system in Norway: A multisite randomized clinical trial in a naturalistic setting. *Psychotherapy Research*, epub ahead of print.

Andrusyna, T. P., Luborsky, L., Pham, T., & Tang, T. Z. (2006). The mechanisms of sudden gains in supportive-expressive therapy for depression. *Psychotherapy Research*, 16, 526-536.

Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). Using client feedback to improve couple therapy outcomes: a randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, 77, 693-704.

Aoun, S., Pennebaker, D., & Janca, A. (2002). Outcome measurement in rural mental health care: A field trial of rooming-in models. *Australian Journal of Rural Health*, 10, 302-307.

Asay, T. P., & Lambert, M. J. (1999). The empirical case for the common factors in therapy: Quantitative findings. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: What works in therapy* (pp. 23-55). Washington, DC: American Psychological Association.

Bankoff, E. A., & Howard, K. I. (1992). The social network of the psychotherapy client and effective psychotherapeutic process. *Journal of Psychotherapy Integration*, 2, 273-294.

Bauer, S. (2004). *Ergebnismonitoring und Feedback: Mittel zur Optimierung stationärer Psychotherapie?* Dissertation zur Erlangung des akademischen Grades Doktor der Sozialwissenschaften in der Fakultät für Sozial- und Verhaltenswissenschaften der Eberhard-Karls-Universität Tübingen.

Berking, M., Orth, U., & Lutz, W. (2006). Wie effektiv sind systematische Rückmeldungen des Therapieverlaufs an den Therapeuten? *Zeitschrift für Klinische Psychologie und Psychotherapie*, 35, 21-29.

Boswell, J. F., Castonguay, L. G., & Wasserman, R. H. (2010). Effects of psychotherapy training and intervention use on session outcome. *Journal of Consulting and Clinical Psychology*, 78, 717-723.

Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2013). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, epub ahead of print.

Boyce, M. B., Browne, J. P., Greenhalgh, J. (2014). The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: A systematic review of qualitative research. *BMJ Quality and Safety*, 23, 508-518.

- Byrne, S., Hooke, G. R., Newnham, E. A., & Page, A. C. (2012). The effects of outcomes monitoring on subsequent readmission to psychiatric care: A six month follow-up. *Journal of Affective Disorders*, 137, 113-116.
- Castonguay, L. G., Barkham, M., Lutz, W., & McAleavey, A. A. (2013). Practice-oriented research: Approaches and application. In M. J. Lambert (Eds.), *Bergin and Garfield's Handbook of psychotherapy and behavior change*, 6th edition (pp. 85-133). New York: Wiley.
- Chambless, D. L., & Hollon S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66, 7-18.
- Chapman, C. L., Burlingame, G. M., Gleave, R., Rees, F., Beecher, M., & Porter G. S. (2012). Clinical prediction in group psychotherapy. *Psychotherapy Research*, 22, 673-681.
- Crits-Christoph, P., Ring-Kurtz, S., Hamilton, J. L., Lambert, M. J., Gallop, R., McClure, B. et al. (2012). A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *Journal of Substance Abuse Treatment*, 42, 301-309.
- Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive psychotherapy for adult depression: A meta-analysis. *Clinical Psychology Review*, 32, 280-291.
- de Jong, K., van Sluis, F., Nugter, M. A., Heisser, W. J., & Spinhoven, F. (2012). Understanding the differential impact of outcome monitoring: therapist variables that moderate feedback effects in a randomized clinical trial. *Psychotherapy Research*, 22, 464-474.
- Dinger, U., & Schauenburg, H. (2010). Effects of individual cohesion and patient interpersonal style on outcome in psychodynamically oriented inpatient group psychotherapy. *Psychotherapy Research*, 20, 22-29.
- Finch, A. E., Lambert, M.J., & Schaalje, B.G. (2001). Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. *Clinical Psychology and Psychotherapy*, 8, 231-242.
- Fischer, F., Schirmer, N., Tritt, K., Klapp, B. F., & Fliege, H. (2011). Retestreliaibilität und Änderungssensitivität des ICD-10–Symptom-Rating (ISR) in verschiedenen Stichproben. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 61, 162-169.
- Fischer, F., Tritt, K., Klapp, B. F., & Fliege, H. (2010). Faktorenstruktur und psychometrische Eigenschaften des ICD-10-Symptom-Rating (ISR) an Stichproben psychosomatischer Patienten. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 60, 307-315.
- Fliege, H., Rose, M., Bronner, E., & Klapp, B. F. (2002). Prädiktoren des Behandlungsergebnisses stationärer psychosomatischer Therapie. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 52, 47-55.

- Frank, J. D., & Frank, J. B. (1991). *Persuasion and healing: A comparative study of psychotherapy, 3rd edition*. Baltimore, MD: Johns Hopkins University.
- Franke, G. H., Hoffmann, T., & Frommer, J. (2005). Entspricht die Symptombesserung vier Wochen nach Behandlungsbeginn dem Erfolg in der Ein-Jahres-Katamnese? *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 51, 360-372.
- Fydrich, T., Geyer, M., Hessel, A., Sommer, G., & Brähler, E. (1999). Fragebogen zur Sozialen Unterstützung (F-SozU): Normierung an einer repräsentativen Stichprobe. *Diagnostica*, 45, 212-216.
- Garnefski, N., Kraaij, V., & Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personality and Individual Differences*, 30, 1311-1327.
- Greenhalgh, J. (2009). The applications of PROs in clinical practice: What are they, do they work, and why? *Quality of Life Research*, 18, 115-123.
- Haase, M., Frommer, J., Franke, G.-H., Hoffmann, T., Schulze-Muetzel, J., Jäger, S. et al (2008). From symptom relief to interpersonal change: Treatment outcome and effectiveness in inpatient psychotherapy. *Psychotherapy Research*, 18, 615-624.
- Haderlie, M. (2009). *Patient-focused research: Examining the psychotherapist as a feedback receiver*. UNLV Theses/Dissertations/Professional Papers/Capstones. Paper 65.
- Hall-Flavin, D. K., Winner, J. G., Allen, J. D., Jordan, J. J., Nesheim, R. S., Snyder, K. A. et al. (2012). Using a pharmacogenomic algorithm to guide the treatment of depression. *Translational Psychiatry*, 2, e172.
- Hall-Flavin, D. K., Winner, J. G., Allen, J. D., Jordan, J. J., Carhart, J. M., Proctor, B. et al. (2013). Utility of integrated pharmacogenomic testing to support the treatment of major depressive disorder in a psychiatric outpatient setting. *Pharmacogenetics and Genomics*, 23, 535-548.
- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, 61, 155-163.
- Hansson, H., Rundberg, J., Österling, A., Öjehagen, A., & Berglund, M. (2013). Intervention with feedback using Outcome Questionnaire 45 (OQ-45) in a Swedish psychiatric outpatient population. A randomized controlled trial. *Nordic Journal of Psychiatry*, 67, 274-281.
- Harbour, R., & Miller, J. (2001). A new system for grading recommendations in evidence based guidelines. *BMJ*, 323, 334-336.
- Harmon, C., Hawkins, E. J., Lambert, M. J., Slade, K., & Whipple, J. S. (2005). Improving outcomes for poorly responding clients: The use of clinical support tools and feedback to clients. *Journal of Clinical Psychology*, 61, 175-185.

- Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. *Psychotherapy Research*, 17, 379-392.
- Hatfield, D., McCullough, L., Frantz, S. H., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology and Psychotherapy*, 17, 25-32.
- Haug, S., Puschner, B., Lambert, M. J., Kordy, H. (2004). Veränderungsmessung in der Psychotherapie mit dem Ergebnisfragebogen (EB-45). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 25, 141-151.
- Hawkins, E. J., Lambert, M. J., Vermeersch, D. A., Slade, K. L., & Tuttle, K. C. (2004). The therapeutic effects of providing patient progress information to therapists and patients. *Psychotherapy Research*, 14, 308-327.
- Henningsen, P., & Rudolf, G. (2000). Zur Bedeutung der evidenced-based medicine für die Psychotherapeutische Medizin. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 50, 366-375.
- Herrmann, A. S., & Huber, D. (2013). Was macht stationäre Psychotherapie erfolgreich? Der Einfluss von Patienten- und Behandlungsmerkmalen auf den Therapieerfolg in der stationären Psychotherapie. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 59, 273-289.
- Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, 48, 9-16.
- Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist*, 41, 159-164.
- Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology*, 61, 678-685.
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, 51, 1059-1064.
- Huber, D., Albrecht, C., Hautum, A., Henrich, G., & Klug, G. (2009). Langzeit-Katamnese zur Effektivität einer stationären psychodynamischen Psychotherapie. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 55, 189-199.
- Ilardi S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavior therapy of depression. *Clinical Psychology - Science and Practice*, 1, 138-155.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.

Jain, K. K. (2002). Personalized medicine. *Current Opinion in Molecular Therapeutics*, 4, 548-558.

Jensen-Doss, A., & Hawley, K. M. (2010). Understanding barriers to evidence-based assessment: Clinician attitudes toward standardized assessment tools. *Journal of Clinical Child and Adolescent Psychology*, 39, 885-896.

Kaufmann, Y., & Weck, F. (2014). *Die Bedeutung von Feedback für die Entwicklung psychotherapeutischer Kompetenzen: Eine randomisierte kontrollierte Studie*. Poster auf dem 32. Symposium Klinische Psychologie und Psychotherapie, der DGPs Fachgruppe Klinische Psychologie und Psychotherapie in Braunschweig.

Knaup, C., Kösters, M., Schöfer, D., Becker, T. & Puschner, B. (2009). Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. *The British Journal of Psychiatry*, 195, 15-22.

Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L.E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, 62, 1009-1016.

Kraft, D., Koellner, V., & Lutz, W. (2008, Juli). *Patientenorientierte Psychotherapieforschung in der stationären Routineversorgung*. Symposiumsbeitrag zur 1. Forschungswerkstatt Stationäre Psychotherapie in Dresden.

Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice*. Washington, DC: American Psychological Association.

Lambert, M. J. (2012). Helping clinicians to use and learn from research-based systems: The OQ-Analyst. *Psychotherapy*, 49, 109-114.

Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, 50, 42-51.

Lambert, M. J., Bailey, R., Kimball, K., Shimokawa, K., Harmon, S. C., & Slade, K. (2007). *Clinical support tools manual-brief version-40*. Salt Lake City: OQMeasures.

Lambert, M. J., Hannöver, W., Nisslmüller, K., Richard, M., & Kordy, H. (2002). Fragebogen zum Ergebnis von Psychotherapie. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 31, 40-47.

Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, 69, 159-172.

Lambert, M. J., Kahler, M., Harmon, S. C., Burlingame, G. M., & Shimokawa, K. (2011). *Administration and scoring manual for the Outcome Questionnaire-45.2*. Salt Lake City: OQMeasures.

Lambert, M. J., Whipple, J.L., Bishop, M. J., Vermeersch, D. A., Gray, G. V., & Finch, A. E. (2002). Comparison of empirically derived and rationally derived methods for identifying clients at risk for treatment failure. *Clinical Psychology and Psychotherapy*, 9, 149-164.

- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice*, 10, 288-301.
- Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology and Psychotherapy*, 9, 91-103.
- Liebherz, S., & Rabung, S. (2013). Wirksamkeit psychotherapeutischer Krankenhausbehandlung im deutschsprachigen Raum: Eine Meta-Analyse. *Psychotherapie, Psychosomatik, medizinische Psychologie*, 63, 355-364.
- Lilienfeld, S. O., Ritschel, L. A., Lynn, S. J., Cautin, R. L., & Latzman, R. D. (2013). Why many clinical psychologists are resistant to evidence-based practice: Root causes and constructive remedies. *Clinical Psychology Review*, 33, 883-900.
- Lutz, W., Martinovich, Z., & Howard, K. I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology*, 67, 571-577.
- Lutz, W., Lambert, M. J., Harmon, C. J., Tschitsaz, A., Schürch, E., & Stulz, N. (2006). The probability of treatment success, failure and duration – what can be learned from empirical data to support decision making in clinical practice? *Clinical Psychology and Psychotherapy*, 13, 223-232.
- Lutz, W., Leach, C., Barkham, M., Lucock, M., Stiles, W. B., Evans, C., Noble, R., & Iveson, S. (2005). Predicting rate and shape of change for individual clients receiving psychological therapy: Using growth curve modeling and nearest neighbor technologies. *Journal of Consulting and Clinical Psychology*, 73, 904-913.
- Maling, M. S., Gurtman, M. B., & Howard, K. I. (1995). The response of interpersonal problems to varying doses of psychotherapy. *Psychotherapy Research*, 5, 63-75.
- Melchior, H. (2011). *Vorhersage des längerfristigen Behandlungserfolgs anhand von Symptomverläufen und der Dauer stationärer Psychotherapie*. Dissertation zur Erlangung der Würde des Doktors der Philosophie am Fachbereich Psychologie, Fakultät für Erziehungswissenschaft, Psychologie und Bewegungswissenschaft der Universität Hamburg.
- Newnham, E. A., Hooke, G. R., & Page, A. C. (2010). Patient monitoring and feedback in psychiatric care reduces depressive symptoms. *Journal of Affective Disorders*, 127, 139-146.
- Okiishi, J., Lambert, M. J., Nielsen, S. L., & Ogles, B. M. (2003). Waiting for supershrink: An empirical analysis of therapist effects. *Clinical Psychology and Psychotherapy*, 10, 361-373.
- Pelletier, L. G., Tuson, K. M., & Haddad, N. K. (1997). Client Motivation for Therapy Scale: A measure of intrinsic motivation, extrinsic motivation, and amotivation for therapy. *Journal of Personality Assessment*, 68, 414-435.

Probst, T., Lambert, M. J., Dahlbender, R. W., Loew, T. H., & Tritt, K. (2014). Providing patient progress feedback and clinical support tools to therapists: Is the therapeutic process of patients on-track to recovery enhanced in psychosomatic in-patient therapy under the conditions of routine practice? *Journal of Psychosomatic Research*, 76, 477-484.

Probst, T., Lambert, M. J., Loew, T. H., Dahlbender, R. W., & Tritt, K. (under re-review). Extreme deviations from expected recovery curves and their associations with therapeutic alliance, social support, motivation, and life events in psychosomatic in-patient therapy. *Psychotherapy Research*

Probst, T., Lambert, M. J., Loew, T. H., Dahlbender, R. W., Göllner, R., & Tritt, K. (2013). Feedback on patient progress and clinical support tools for therapists: improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice. *Journal of Psychosomatic Research*, 75, 255-261.

Probst, T., Loew, T. H., Dahlbender, R. W., & Tritt, K. (2014). Belastungsgrade zur Aufnahmewoche und wöchentliche Änderungsmuster in der stationären Psychosomatik. *Die Psychodynamische Psychotherapie*, 13, 75-90.

Probst, T., von Heymann, F., Zaudig, M., Konerman, J., Lahmann, C. Loew, T., & Tritt K. (2009). Effektivität stationärer psychosomatischer Krankenhausbehandlung - Ergebnisse einer multizentrischen Katamnese studie. *Zeitschrift für Psychosomatische Medizin*, 55, 409-420.

Puschner, B., Schöfer, D., Knaup, C. & Becker, T. (2009). Outcome management in in-patient psychiatric care. *Acta Psychiatrica Scandinavica*, 120, 308-319.

Reese, R.J., Usher, E.L., Bowman, D.C., Norsworthy, L.A., Halstead, J.L., Rowlands, S.R. et al. (2009). Using client feedback in psychotherapy training: An analysis of its influence on supervision and counselor self-efficacy. *Training and Education in Professional Psychology*, 3, 157-168.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *BMJ*, 312, 71-72.

Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology*, 61, 145-153.

Schneider, W., Klauer, T., Janssen, P. L., & Tetzlaff, M. (1999). Zum Einfluß der Psychotherapiemotivation den Psychotherapieverlauf. *Nervenarzt*, 70, 240-249.

Sexton, H. (1996). Process, life events, and symptomatic change in brief eclectic psychotherapy. *Journal of Consulting and Clinical Psychology*, 64, 1358-1365.

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78, 298-311.

Simon, W., Lambert, M. J., Busath, G., Vazquez, A., Berkeljon, A., Hyer, K. et al. (2013). Effects of providing patient progress feedback and clinical support tools to psychotherapists

in an inpatient eating disorders treatment program: A randomized controlled study. *Psychotherapy Research*, 23, 287-300.

Simon, W., Lambert, M. J., Harris, M. W., Busath, G., & Vazquez, A. (2012). Providing patient progress information and clinical support tools to therapists: Effects on patients at risk of treatment failure. *Psychotherapy Research*, 22, 638-647.

Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology and Psychotherapy*, 15, 287-303.

Spielmans, G. I., Masters, K. S., & Lambert, M. J. (2006). A comparison of rational versus empirical methods in prediction of negative psychotherapy outcome. *Clinical Psychology and Psychotherapy*, 13, 202-214.

Spring, B. (2007). Evidence-based practice in clinical psychology: What it is; why it matters; what you need to know. *Journal of Clinical Psychology*, 63, 611-631.

Steffanowski, A., Löschmann, C., Schmidt, J., Wittmann, W.W. & Nübling, R. (2007). *Metaanalyse der Effekte stationärer psychosomatischer Rehabilitation*. Bern: Huber.

Strauss, B., Kirchmann, H., Eckert, J., Lobo-Drost, A., Marquet, A., Papenhausen, R. et al. (2006). Attachment characteristics and treatment outcome following inpatient psychotherapy: Results of a multisite study. *Psychotherapy Research*, 16, 579-594.

Stulz, N., Lutz, W., Kopta, S.M., Minami, T., Saunders, S. M. (2013). The shape of the dose-effect relationship in routine outpatient psychotherapy: Does treatment duration matter? *Journal of Counseling Psychology*, 60, 593-600.

Tang, T. Z., & DeRubeis, R.J. (1999) Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 67, 894-904.

Tang, T. Z., DeRubeis, R. J., Beberman, R., & Pham, T. (2005). Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 73, 168-172.

Tschitsaz-Stucki, A., & Lutz, W. (2009). Identifikation und Aufklärung von Veränderungssprüngen im individuellen Psychotherapieverlauf. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 38, 13-23.

Tritt, K., von Heymann, F., Zaudig, M., Loew, T., Söllner, W., Fischer, H. F., & Bühner, M. (2010). Das ICD-10-Symptom-Rating (ISR): Zusammenhänge mit der SCL-90-R. *Zeitschrift für Klinische Diagnostik und Evaluation*, 3, 22-37.

Tritt, K., von Heymann, F., Zaudig, M., Probst, T., Loew, T., Klapp, B., Söllner, W., & Bühner, M. (in Vorbereitung). *Der Fragebogen ICD-10 Symptom-Rating (ISR) – das Manual*.

Tritt, K., von Heymann, F., Zaudig, M., Zacharias, I., Söllner, W., & Loew, T. (2008). Entwicklung des Fragebogens „ICD-10-Symptom-Rating“ (ISR). *Zeitschrift für Psychosomatische Medizin*, 54, 420-429.

Tschuschke, V. (2005). Die Psychotherapie in Zeiten evidenzbasierter Medizin. *Psychotherapeutenjournal*, 106-115.

von Heymann, F., Zaudig, M., Tritt, K. et al (2003). Die diagnosebezogene Behandlungsdauer in der Psychosomatischen und Psychotherapeutischen Medizin: Eine homogene Größe? Erste Ergebnisse der Multicenter-Basisdokumentation (Psy-BaDo-PTM) als Grundlage qualitätssichernder Maßnahmen in der stationären Psychosomatik. *Praxis Klinische Verhaltensmedizin und Rehabilitation*, 62, 209–221.

Walter, W. G., Cleary, M., & Rey, J. M. (1998). Attitudes of mental health personnel towards rating outcome. *Journal of Quality in Clinical Practice*, 18, 109-115.

Whipple, J. L., Lambert, M. J., Vermeersch, D.A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology*, 50, 59-68.

Wilmers, F., Munder, T., Leonhart, R., Herzog, T., Plassmann, R., Barth, J., & Linster W. (2008). Die deutschsprachige Version des Working Alliance Inventory – short revised (WAI-SR) – Ein schulenübergreifendes, ökonomisches und empirisch validiertes Instrument zur Erfassung der therapeutischen Allianz. *Klinische Diagnostik und Evaluation*, 1, 343-358.

Wittmann, W. W., Lutz, W., Steffanowski, A., Kriz, D., Glahn, E. M., Völkle, M. C., Böhnke, J. R., Köck, K., & Bittermann, A. (2011). *Qualitätsmonitoring in der ambulanten Psychotherapie: Modellprojekt der Techniker Krankenkasse - Abschlussbericht*. Techniker Krankenkasse.